DOCUMENT RESUME

ED 177 928 HE 011 313

TITLE	Assessment at Alverno College.
INSTITUTION	Alverno Coll., Milwaukee, Wis.
SPONS AGENCY	Fund for the Improvement of Postsecondary Education (DHEW), Washington, D.C.; Kellogg Foundation, Battle Creek, Mich.
PUB DATE	79
NOTE	65p.
AVAILABLE FROM	Alverno College, 3401 S. 39th Street, Milwaukee, WI 53215
EDRS PRICE	MF01/PC03 Plus Postage.
DESCRIPTORS	Church Related Colleges; College Students; *Competence; *Evaluation Methods; Higher Education; *Liberal Arts; Single Sex Colleges; Small Colleges; *Student Development; *Student Evaluation
IDENTIFIERS	*Alverno College WI; *Outcomes of Education

ABSTRACT

The Alverno College faculty reports on its development and use of assessment (rather than testing) to elicit, diagnose, and certify student abilities as part of its outcome-centered approach to liberal education. Assessment is defined as a more complete challenge to both the process and the content of a student's learning, achieving a more personal, specific and integrated view of her developing competence. The book raises several inevitable questions (e.g., reliable observing, accurate sampling of a student's abilities, defining appropriate criteria), and explains how the faculty of this small, urban Catholic college for women discovered assessment as essential to their approach to liberal learning. It is shown that the authors' view of competence as "generic," "developmental" and "holistic" requires certain qualities in each Alverno assessment technique. Ways in which faculty members design assessments are demonstrated, as well as how assessors are recruited and trained. In one chapter, students reflect on the experience of being assessed and learning to assess. A final chapter reviews the elements of the college's assessment system, and assessment's impact on the faculty. The book repeatedly stresses beginning with a clear focus on the outcomes to be assessed, in developing individual techniques or an overall system. It is emphasized that assessment is a learning experience for students and faculty alike. (Author/LBH)

****
* Reproductions supplied by EDRS are the best that can be made *
* from the original document. *
****

ED177928

Assessment at Alverno College

"PERMISSION TO REPRODUCE THIS MATERIAL HAS BEEN GRANTED BY

Alverno College

TO THE EDUCATIONAL RESOURCES INFORMATION CENTER (ERIC)."

U S DEPARTMENT OF HE... EDUCATION & WELF... NATIONAL INSTITUTE EDUCATION

THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

HE 011 313

Assessment at Alverno College

By the Alverno College Faculty

The program described here was funded in part by the Fund for the Improvement of Postsecondary Education, U.S. Department of Health, Education, and Welfare. The publication of this report was made possible through a grant from the Kellogg Foundation.

The opinions are those of the faculty of Alverno College. No endorsement by either the Fund or the Kellogg Foundation should be inferred.

Copyright © 1979 by Alverno Productions.

Table of Contents

I. How We Define and Why We Do Assessment ... 1

II. How We Understand Assessment 7

III. How We Design and Do Assessment 13

IV. How Students Experience Assessment 33

V. How We Develop Our Assessment System 51

Alverno College

3401 S. 39th St. / Milwaukee, WI 53215

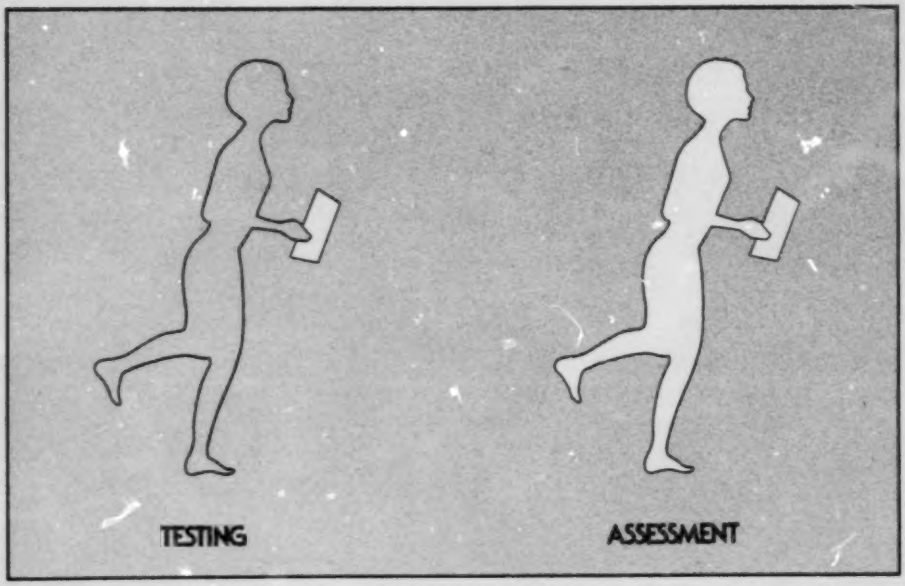
I. How We Define and Why We Do Assessment

To assess. This term, which at its simplest means to observe and to judge something, is one of the most visible results of what has been called a "Copernican revolution" in education.*

Today, most American educators would probably agree that the learner — not the teacher — is at the center of the educational universe. In the last few decades, however, growing numbers of educators have seriously set about translating the modern consensus into action. In varying ways, they have undertaken the task of helping learners learn certain processes — how to seek out, integrate and use knowledge — rather than simply passing along the body of knowledge itself.

Education, so understood, can never be adequately evaluated by traditional testing. Narrow, one-dimensional probes into a student's mines of stored information do not begin to get at how she learns or what she can do. Grades and curves, which sort students into groups for administrative handling, say nothing about how each one is using her talents or growing toward her potential.

This new view of learning demands a broader yet more personal view of each learner's progress. Hence, assessment: a multidimensional attempt to observe and judge the individual learner in action.

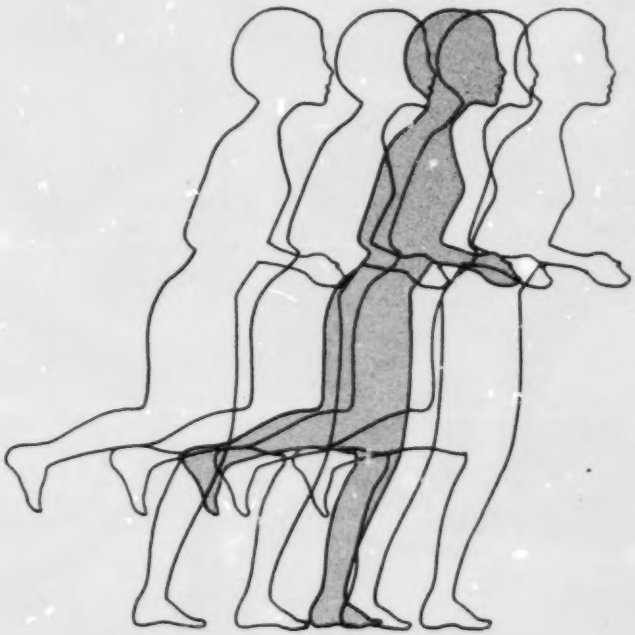


*Russell Edgerton, From Testing to Assessment; Fund for the Improvement of Post-Secondary Education, Washington, D.C., 1976.

There are problems. To begin with, we must rely on *sampling*. A human being is a living mystery, only partly visible to another person. Yet even that part would take more than a lifetime to record. From the tens of thousands of hours of a student's college career, we can select but a handful for careful observation.

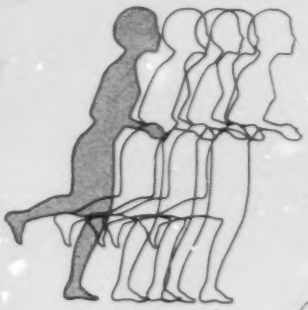
How do we know the sample is representative? Is the student having an unusually bad (or good) day? Is she simply showing her strong suit, or can she exhibit the same abilities under different circumstances?

Then there is *observing*. Even the smallest sample of human behavior is so richly complex that we can focus only on a few key elements. Which ones? How do we record them? Does the situation really call for them? Does it restrict the student to doing only what we are looking for, or does it challenge her to show her best?



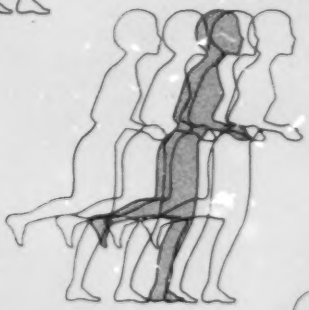
SAMPLING

Judging raises more hard questions. How can we separate an intangible ability (like analysis) from the activity (like writing) that makes it visible? Who shall judge? How do we assure some uniformity among several assessors? In judging we are *applying criteria*. How much demonstration of an ability is enough? What makes an "effective" action? How are these and other criteria set? By whom?

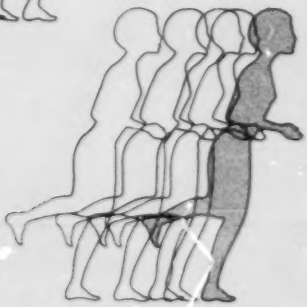


These questions are not the kind that have fixed answers. To undertake assessment means sampling, observing and judging according to criteria. It means developing a whole array of techniques, to take into account the fullest possible range of human talents. And it means an ongoing commitment to dealing with these kinds of questions.

The following pages record some of our exploration into assessment at Alverno College. We have accepted the challenge because it is full of promise. Already we are learning to ask more and better questions. We are developing a more solid ground for comparisons and evaluations, since assessment is built on specified criteria. And we are gaining more and more individualized portraits of each learner's unique constellation of unfolding abilities, her personal style of learning and doing.



CRITERIA



We have made a major commitment to assessment at Alverno College. We are committed to assess what is being taught in every course in our curriculum. Indeed, we consider the assessment techniques and criteria as important as the topics and texts of the courses. We are also committed to assessing student competence outside the classroom. The Assessment Center, with a full-time staff and dozens of trained campus and outside assessors, administers external assessments and evaluative feedback to students throughout the school year.

This all came about as a direct result of our effort to define and bring into being a fundamentally new approach to the age-old task of liberal education. We began this effort a decade ago, in the late 1960s, when serious questions were surfacing nationwide about the meaning and value of college and of liberal education in particular. At Alverno, these questions combined with our need to redefine our mission as a small, urban Catholic liberal arts college for women. The result was a deep and serious inquiry into our goals and values as educators.

After two years of special seminars, faculty institutes, student-faculty commissions and college-wide forums, our inquiry took more definite shape in a series of questions our president asked the academic departments in 1970-71:

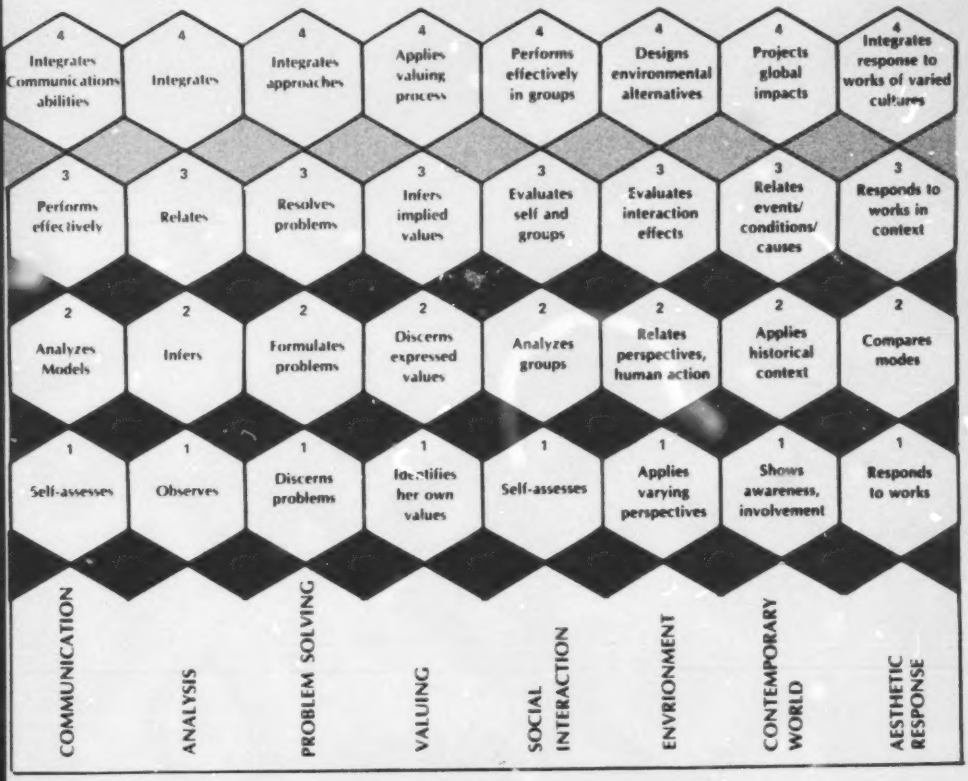
- **"What kinds of questions are being asked by professionals in your field that relate to the validity of your discipline in a total college program?"**
- **"What is your department's position on these?"**
- **"How are you dealing with these problems in your general education courses, and in the work for a major in your field?"**
- **"What are you teaching that is so important that students cannot afford to pass up courses in your department?"**

(The first four levels of each competence, required of every student in her general education program, are detailed in the diagram.)

During the ensuing year of rationale presentations by each of the departments, the faculty reached a consensus that outcomes for the student are the demonstrable value of any learning experience. "What kind of person," we then asked "are we as educators seeking to develop? What outcomes or characteristics will she need as a part of her life?" In response to this question we developed the framework of a series of eight general abilities which, taken together, would result from a successful liberal education.

There are:

1. **Effective communications ability**
2. **Analytical capability**
3. **Problem solving ability**
4. **Valuing in a decision making context**
5. **Effective social interaction**
6. **Effectiveness in individual/environment relationships**
7. **Responsible involvement in the contemporary world**
8. **Aesthetic responsiveness**



No one of these, we realized, could be taught or learned directly in a single experience. Nor could they be divorced from the liberal arts curriculum of which they are the outcomes. The faculty therefore eventually analyzed each of these abilities — which we called "competences" — into a sequence of six levels at which the student would be expected to demonstrate her ability as she progressed through the undergraduate curriculum.

We identified these levels by examining the existing curriculum in each of our disciplines. Traditionally, each department had described its curriculum as a structure of *knowledge*, beginning with basic general concepts and progressing toward more complex and specialized studies. This time, we worked from the assumption that there is also a progression of *abilities* implicit in the movement from introductory survey to advanced seminar. Our focus, then, was to discern the developmental patterns already embedded in the normal curriculum of our disciplines, rather than to redefine our fields or to create a whole new curricular structure.

After we had identified these competences and their sequential levels, the question became: "How can we tell how far along a student is in developing these competences?" It would be pointless, we realized, to have spent so much time and effort articulating our educational goals unless we were willing to make a similar investment in assessing their attainment. Thus we arrived at assessment as a natural and crucial part of our approach to liberal education.

II. How We Understand Assessment

Once we had come to see our purpose as helping each student develop her abilities, we needed to ask not simply "What does she know?" but "What can she do with what she knows?" This led to more difficult questions than the traditional paper-and-pencil testing methods could probe. No longer could we rely on seeking right answers, or asking for repetitions and reformulations of memorized information. Our assessment techniques would have to challenge each student to show her developing abilities in the best and broadest manner of which she was capable.

Our outcome-centered view of education also considerably deepened our understanding of the two classic functions of testing — to credential and to diagnose. We would need assessments reliable enough for us to stand accountable on them before other educators, both within and outside our college. We also would need assessments sufficiently specific and detailed for the student and her instructors to plan her further learning experiences effectively. Yet we could never afford to lose sight of the ultimate *raison d'etre* for assessment: to provide the student, at each of many steps in her development, with progressively fuller and more individual profiles of her emerging combination of gifts, skills and styles, so that she can become an independent learner.

Finally, we had come to focus on outcomes by acknowledging that learning's value lies in its enduring impact on the learner. This reflected our long-standing intention to foster habits of mind and action that would shape and enhance our students' lives and environments, far beyond the college experience. Our assessments, then, would have to provide some assurance that the abilities a student showed were at some point becoming personalized, part of her habitual way of handling situations in her personal and social as well as her academic and professional life.

As these understandings grew, we took hold of three terms that helped to summarize and guide what we were learning. Our assessments had to be, as nearly as we could make them, generic, developmental and holistic.

Generic

The student competences themselves are generic, rather than simply specific to a task or situation. They are attributes or characteristics of the student herself. While they are learned and demonstrated in specific settings, these generic abilities can be transferred and modified to a variety of situations.

Because of the generic nature of the abilities we are teaching for, therefore, our assessments can and must be both **external** and **multiple**. Both these qualities help to insure that the ability a student develops in a given learning situation is a part of her personal repertoire, and that she can exercise it in varying contexts.

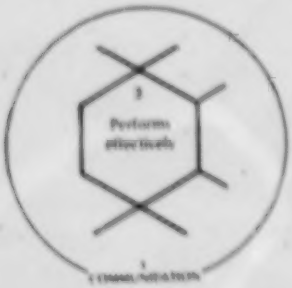
Externality appears in several ways. The criteria used to assess for a given competence level are developed outside the particular course setting and adapted for use in the actual assessment. Assessors often are drawn from outside the class and off campus. When the student and instructor perform as assessors, they do so with a certain role distance from their initial involvement in the exercise.

Multiplicity occurs in requiring a single ability to be validated in several settings, and in the constant use of multiple viewpoints — instructors and faculty, co-assessors, off-campus assessors, peers and self-assessment. The use of multiple assessments gives the student the opportunity and the obligation to seek certification in a variety of settings. In addition, she always has at least two sources of judgment (herself and another assessor) and often several, from a variety of backgrounds.

Thus, for example, one part of effective communications is the ability to deliver a complex spoken message to an uninitiated audience. Attempting to demonstrate this ability, a student might well prepare and deliver a talk on cardiovascular circulation for an eighth-grade class. (Such an exercise, we have found, provides a powerful test of how well she has "made her own" the physiology of the heart and blood vessels, along with a sample of her speaking abilities in action.) Her physiology instructor and a communications instructor would join in judging such specific criteria as organization (thesis statement, supporting evidence), voice projection, and so on. Later, the student would receive feedback on the strengths and successes in her talk as well as the areas in which she next needs to concentrate her learning effort.

But before the faculty can assert that she has made this ability a usable part of her personal communication repertoire, a successful assessment (formally referred to at Alverno as a "validation") must be corroborated. She would also be required to deliver comparable complex spoken messages successfully in other settings — say, before a simulated urban planning group and in a philosophy seminar.

The assessments used to establish a given level of competence may vary widely, both in the modes the student employs and in the methods of assessment. Demonstrating her ability at a given level in analysis, for example, a student might find herself called upon to draft and later edit a written critique of the sound and image patterns in a poem, with her English instructor and herself as assessors. She might also work with a lab partner to isolate and identify a particular cell variety under a microscope, with herself and her partner, another lab team and finally the biology instructor assessing the lab journal. At a later level, she might participate in an off-campus project identifying patterns in labor grievances and analyzing their underlying causes for a large manufacturing firm, with the company's labor relations director and her management instructor cooperating as assessors.



Developmental

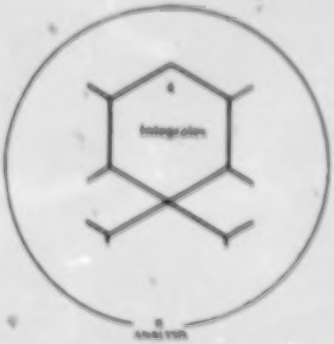
We also understand competence in developmental terms. This means not only that competence levels are sequenced in a progressive learning pattern, but that we use assessment itself as a teaching tool. We do this by making assessment techniques and criteria **public and explicit** and by presenting the results to the student immediately and in detail in a **structured feedback** situation.

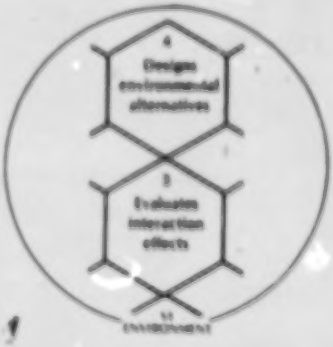
At the beginning of each course, the instructor spells out the course goals in the syllabus, including the competence levels a student may be ready to demonstrate as a result of the course experience. In the syllabus, the instructor relates the particular materials and assignments to the course goals. The syllabus also specifies the assessment techniques that will be used, and enumerates the criteria upon which the students' work will be judged.

In an environmental psychology class, for example, students may seek validation in analysis at level 4 ("integrates patterns seen in data to form conclusions") and in individual/environment relationships at levels 3 and 4 ("specifies the effects of a setting on human behavior" and "designs alternate settings to solve an environment problem"). In weekly field observations each student is required to record a given site's influence on such variables as privacy or stress. This challenges her ability to identify environmental effects (environment, level 3), while required readings provide her with models for integrating what she observes and infers (analysis, level 4).

The student must report orally on her field observations each week. Her instructor and fellow students assess her report, using such criteria as specific description of the setting, of the subjects, and of the behaviors observed, and clear articulation of how she inferred the setting's impact. Besides these weekly assessments, the student is also given larger, more formal assessment problems. Having read and discussed Edward Hall's The Hidden Dimension, for instance, she is asked to design a cafeteria for the UN building, taking into account the wide differences Hall outlines in cultural space/intimacy patterns. This exercise challenges her ability to propose environmental alternatives (environment, level 4).

Whether or not she elects to attempt validation during the course, each student receives ongoing assessment. This gives both her and her instructor a diagnostic view of her learning progress. As soon as possible after each assessment, the student receives detailed feedback on her performance in terms of the overall course goals and specific criteria set forth in the syllabus.





In the environmental psychology course, feedback on the field work assessments is instantaneous (though the instructor may also set up an individual feedback session with a student he feels is having difficulties). On exercises like the UN cafeteria design, the student might receive a detailed written annotation of her plan and her diagrams, and the class might engage in a group feedback session as well. The feedback may also address skills the student exhibits in collateral areas (such as graphing, aesthetics or social interaction in her UN cafeteria design), even though they are not being validated in this course.

Whether it is for validation or for diagnosis alone, whether it is as simple as a series of one-paragraph responses to questions about a film or as complex as presenting a park use plan to a neighborhood association, we try to use each assessment situation as a learning experience. Ideally, assessment should contribute to and culminate a process of working toward explicit, known goals, with frequent stops to measure "the state of the art" in the ability the student is working to develop and we are working to foster.

As opposed to most traditional testing, no element of surprise or secrecy is needed in assessment. In fact, we find that developmental growth flourishes where the goals and purposes of teaching, along with the techniques for pursuing them and measuring their achievement, are made explicit and public. Growth is further enhanced when the results of assessment are presented in structured feedback experiences in which the student can expect a careful commentary on her performance, its elements and its implications.

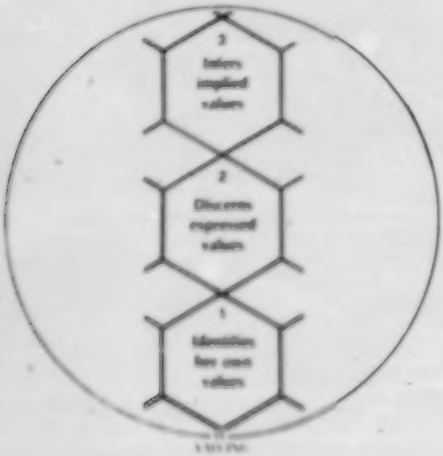
Holistic

The third defining concept we recognized as we developed our assessment process is the holistic nature of the competences. We have analyzed them into levels and component abilities in order to make them manageable for effective teaching, learning and assessment. Yet we realize that the competences are, in reality, inseparable parts of the whole person. Taken together, they answer the first question we asked ourselves about outcomes: "What kind of person, with what characteristics, are we seeking to develop?"

The sequential levels of the competences are therefore **cumulative**, in both experience and assessment. Each level builds upon and includes the prior levels, and the student has frequent opportunity to further evaluate and refine abilities for which she has already been validated. Indeed, we constantly coordinate our instructional efforts to ensure that this automatically occurs. As she undertakes to specify increasingly complex relationships among world events (contemporary world, level 3), for example, she must begin interweaving the separate historical backgrounds they arose from — a sophistication of her proven ability to explore a single event's historical context (level 2). Prior levels of ability are thus not only reinforced, but are drawn into more and more complex uses.

At the same time, learning and assessment look forward beyond the level at hand. In each course or individual learning experience, and in each assessment situation, we attempt to elicit from the student the most advanced performance of which she is capable. For example, while keeping a journal to find and make explicit what her values are (level 1 of valuing), a student is likely to explore her actions, her aesthetic preferences or her religious heritage. She might record reflections on various products of our technological age, from modern medicine to transistor radios.

Her instructor would examine the journal according to the criteria for level 1. In addition, the instructor would point out where the student discerns expressed values in artistic and cultural works (level 2) or infers value problems implicit in technology (level 3). The feedback would thus show her where she is already developing abilities beyond those for which she is seeking validation. The **unlimited** possibilities for exhibiting competence are even more



evident at the advanced levels. Here the student helps design or select highly complex assessment situations, often in professional or other field settings, which challenge her abilities to their fullest.

The criteria by which the student's performance is judged are necessarily specific to the level for which she is seeking validation. But with precise observation as a basis, the assessor can move in feedback beyond the level at hand to discuss the broader implications of the student's performance. In recording exact information about the student's performance, assessors may be called on to judge whether a given element is present or absent, to quantify its frequency or to use a rating scale. Always, they are expected to note illustrative examples. In both qualitative and quantitative measures, however, we take care not to place a ceiling on the degree to which a student may excel in demonstrating the criteria involved. Indeed, our constant emphasis in teaching and assessing is to encourage the student to go as far as she can with her developing abilities at every opportunity.

Finally, the competences and their assessment are by nature **integrated**. This, too, is most apparent at the advanced levels. Early assessments are fairly well enclosed in their design, in order to focus the student's attention (both as a performer and as an assessor) on a particular element of a single ability. By the time she is working toward advanced level validations, however, we are looking explicitly at several competences operating in concert.

When a nursing student is ready to attempt level 5 in problem solving, she invites an instructor into her field setting for a clinical co-assessment. She may choose to seek validation for social interaction and valuing at the same time. Her written "nursing care plan" for each client records her command of problem solving as applied to nursing. As she deals with her clients face to face, she carries out her problem solving in action and also displays her interactive abilities. In diagnosing a client's needs, assigning them priorities, and evaluating her plan and its effectiveness, she also applies her abilities in valuing. From the student's nursing care plans, the co-assessor's careful record of her clinical performance and the student's own critique of her work, she can be assessed in all three competences: social interaction, valuing and problem solving.

In like manner, an English student seeking validation at levels 5 and 6 would be given a week-long simulation exercise. At this level she must be able to analyze, respond to and evaluate complex literature that represents a variety of writers and historical periods. She must also be able to extend to human experience in general the understanding of multiplicity of point of view that she has developed in the context of literature.

Working two to three hours a day with several peers as the staff of a fictitious community cultural center, she handles a variety of problems. While planning an upcoming literary festival, she might be asked to step in as emergency substitute teacher in an adult class on Elizabethan plays. She may have to deal personally (on videotape) and in writing with a benefactor's repeated attempts to influence the poetry selections for the festival. On short notice, she may be sent to appear on a radio talk show (also videotaped) to respond to citizens of varied perspectives who call in to criticize city plans for razing a block of tenements in order to expand the center.

In a variety of ways, she is called on to apply her literary knowledge, her ability to define and defend criteria for judging works, and her understanding of the impact of literary art on its audience (aesthetic response). At the same time, she must frame and deliver complex messages to varying audiences using several media (communications). And she must repeatedly draw together a variety of works in terms of their commonalities and contrasts (analysis). The written, sound taped and videotaped records of her week's work thus provide an ample basis for assessing the student's abilities in three different competence areas.

Beyond enhancing the immediate value of our assessment techniques, our insistence on a holistic focus has helped to remind us of the surprising variety of ways in which a given ability can be demonstrated. Our students' myriad variations upon the theme set by a give instrument continually broaden our awareness of what to look for, as well as providing concrete evidence of their own unique, developing styles. We are thus enabled, as assessors, to appreciate a wide range of possible definitions for "successful performance." We also learn more about how to design our assessment experiences, fine tuning them for specificity and opening them to elicit the richest possible response.

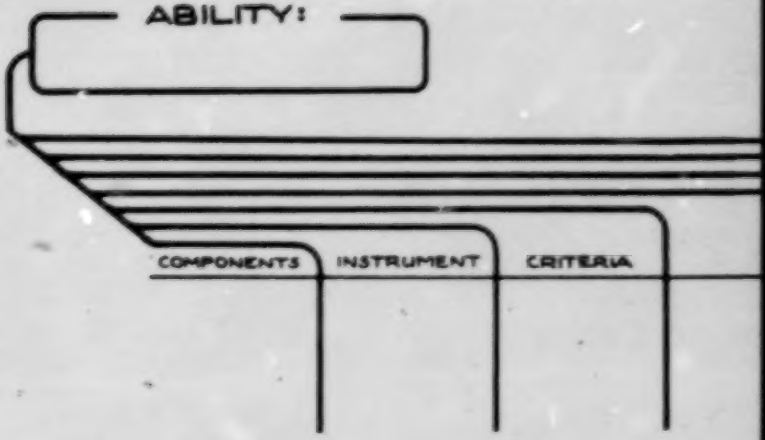
III. How We Design and Do Assessment

When we began designing our first assessment techniques five years ago, we did not rely on a proven method or manual — there was none to be had. Nor did we start by trying to write one. We began with individual faculty members and departments struggling to take abilities that were implicit as outcomes of their courses and turn them into something concrete and observable that students could be expected to do.

By now we have designed, used and modified or discarded literally hundreds of separate assessment techniques. In the process, we have explored countless approaches to the problem of how to design an assessment. Finding a single authorized method was not our goal; we have found instead a principle and a pattern that seem to underlie all our successful approaches.

The principle is simple: to *begin by being as clear as possible about the outcomes*. We could not begin to work out the implications of our outcome-centered view of education until we became specific and detailed about *which* outcomes we had in mind and *how* they could be understood developmentally. Likewise, we have found that we cannot begin to create an effective assessment situation until we are clear about *which* specific abilities we are looking to see demonstrated and *how* they might appear. And we cannot hope to settle the details of our design, nor to observe and judge the student's performance, until we settle the question of *which* criteria we will use and *how* we will apply them. At each succeeding level of specificity, then, from the overall curriculum to assessing a single student performance, we start by seeking the clearest possible view of the outcomes.

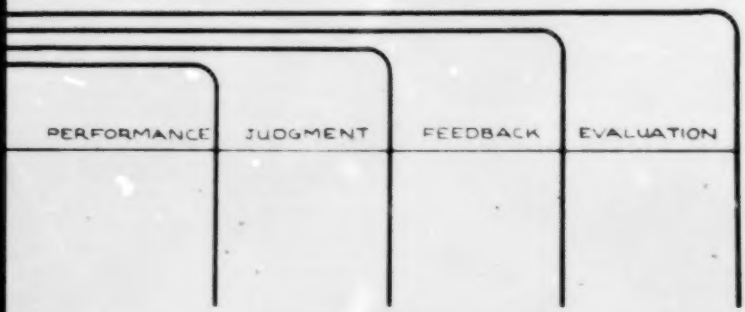
The pattern is also simple. It flows from the fact that we begin with outcomes. While it resembles a sequence of steps (as in the diagram below) it is not a definitive formula nor a fixed procedure. We use it as a "heuristic," a model to stimulate and extend creative thinking. We have found this heuristic valuable not as a sure-fire recipe or checklist, but as a way to raise challenging questions. Whatever ideas or problems we may have in mind as we begin designing a particular assessment technique, we can often use this model to help draw out their implications and bring the various elements of the assessment into sharper focus.



A heuristic is an alternative to trial and error. It is simply the codification of a useful technique or cognitive skill. It can operate as a discovery procedure or a way of getting to a goal. Many fields have them; for example, the scientific method is itself a heuristic, as is journalism's efficient Who? What? When? Where? Why? formula for collecting information. The important thing about heuristics is that they are not rules, which dictate a right or wrong way, but are alternative methods for doing something – methods which often formalize the efficient procedure a good scientist or journalist would use unconsciously. Because they make an intuitive method explicit, heuristics open complex processes up to the possibility of rational choice.

Linda S. Flower
John R. Hayes

College English
December, 1977

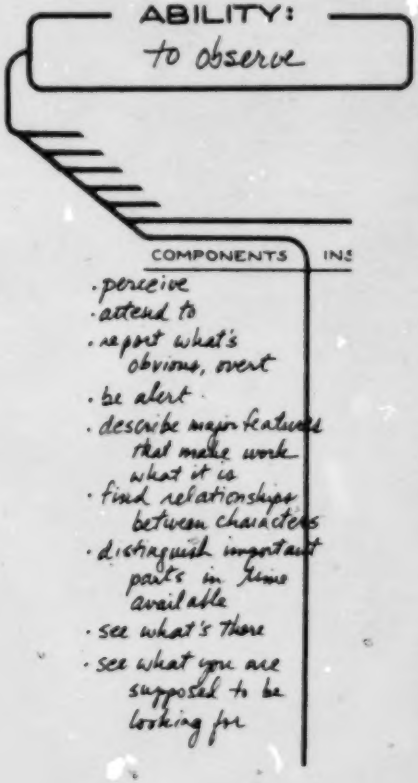


Using the Heuristic with a New Design

An instructor planning an introductory literature course, for example, might well choose to offer analysis as a competence in which her students may seek validation. In the poetry section, she knows she can expect the classic problem that arises when one or two overzealous students rush off to hunt for "hidden meanings" while others who "just can't see all that in a poem" conclude that they had better forget about poetry altogether. To deal with this problem, she might wish to get the class into the habit of looking carefully at a poem before asking anything at all about its meaning.

Such a goal fits in well with the faculty's general definition of the first level of analysis, "to observe." This would be an appropriate level to offer, then; and the instructor now has an assessment to design.

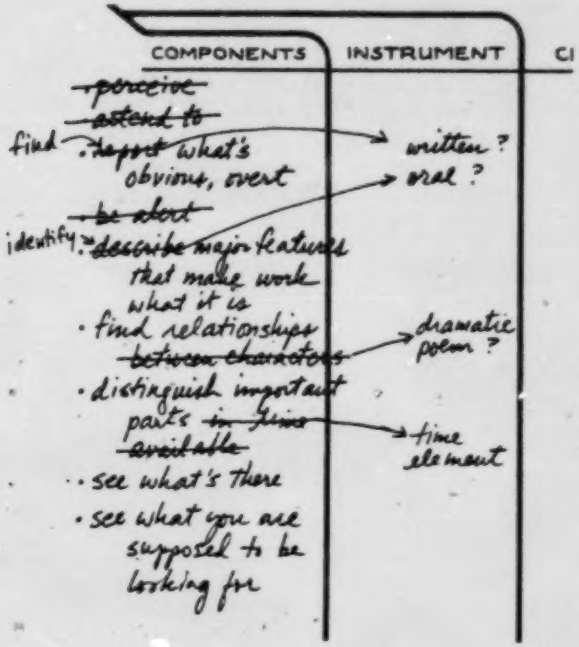
Her work would begin with the outcomes, trying to break out the various elements of successful observing. What does it *mean* to observe? What does a person who is observing a poem well *do*? On her own, or brainstorming with colleagues, the instructor might well arrive at a list something like this:


ABILITY:
to observe
COMPONENTS
INS
· perceive
· attend to
· report what's obvious, overt
· be alert
· describe major features that make work what it is
· find relationships between characters
· distinguish important parts in scene available
· see what's there
· see what you are supposed to be looking for

Once this open generating of ideas seems fairly well complete, the problem becomes one of sorting and ordering. Some of these supposed elements, for example, are really synonyms. "Perceive" and "attend to" restate the whole concept rather than breaking it down into more usable parts. These can be set aside (though not discarded — they may be quite helpful later).

Other items on the list are really conditions for the performance of observing. "Be alert," for instance, is a precondition of any good observation, and may simply have to be inferred from the more directly visible parts of the performance. "Find relationships *between the characters*" includes an assumption about the design of the assessment situation, as does "distinguish important parts *in the time available*." The first assumes that a narrative or dramatic poem will be used, while the second draws attention to the fact that some sort of time limit will need to be set.

Other implied conditions can be seen in "*describe* the major features that make the work what it is" and "*report* what's obvious, what's overt," both of which point toward the need to make the student's observing visible to the assessor. These and the other conditions can be separated out and noted for later use in constructing the details of the assessment situation.



The items that remain, then, begin to look much more like actual components of the ability to observe. Some of the ideas may still need to be broken open. "Find what's obvious, what's overt," for example, clearly suggests that some of what the student might be inclined to report is not obvious or overt — that it is something she has inferred. The instructor might therefore restate this as "distinguish what's obvious, overt, from what's implied, covert."

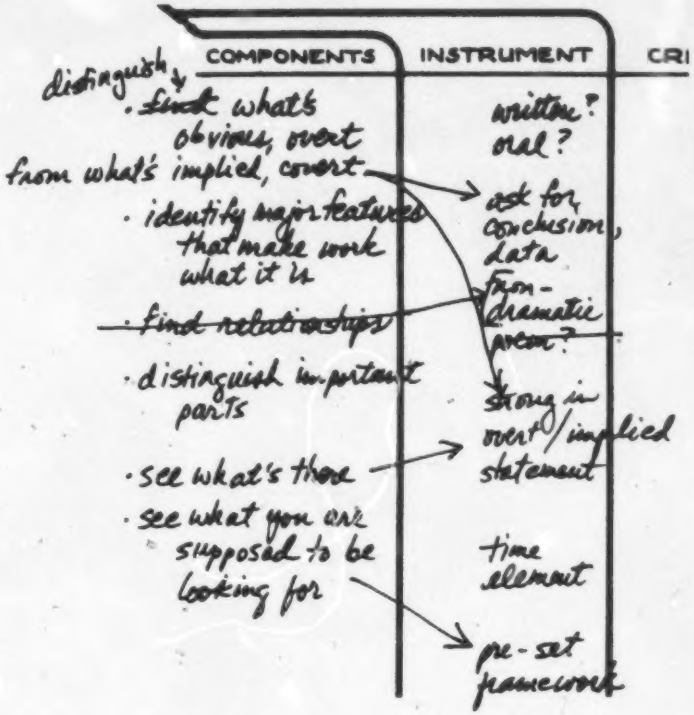
In setting priorities among these several components, the instructor begins moving into the design of the instrument as well. "See what you are supposed to be looking for," for instance, points toward the inevitable need for some kind of framework in observing. Are we looking at a poem as a stream of words to be counted and sorted grammatically? A black-and-white graphic composition? A typographic work? A series of images? At some point, the student must come to realize that frameworks are always involved in observation and to recognize which frameworks she is employing. At this level, however, the instructor may decide not to make that realization a specific goal of instruction. In designing the assessment situation, then, she can control or pre-set this variable by providing the framework in her instructions to the student.

The instructor may likewise decide that "find relationships" represents too complicated a process at this point. Her concern, after all, is to help students develop the habit of doing the simplest kind of observing first. (Since relationships will be the explicit focus of the student's learning experiences at level 3 of analysis, the instructor can set this ability aside knowing it will receive ample attention.) She will also probably decide against a dramatic poem, since the interactions between the characters almost force the reader to focus on relationships.

"See what's there" and "distinguish what's obvious, overt, from what's implied, covert" will be high on her priority list in this assessment. In fact, she may decide to help focus attention on the difference between "what's there" and what the student infers by asking her to draw a conclusion and to cite the overt information that suggested it. The instructor will therefore be obliged, in setting up the problem or "stimulus" to which the student will be asked to respond, to use a poem in which the author both plainly states some things and deliberately implies others.

Let us assume that the instructor has planned to include several poems by Robert Frost. From these she might set aside "In Hardwood Groves," a clear instance of overt and implied statement, for use in the assessment.

Her instrument design, then, could consist of a single class hour in which each student is given a copy of the poem and a set of instructions for a two-part exercise. The first part might ask the student to spend 30 minutes reading the poem and writing an exact account of the things and events


COMPONENTS
INSTRUMENT
CRI
distinguish
· find what's obvious, overt from what's implied, covert
· identify major features that make work what it is
· find relationships
· distinguish important parts
· see what's there
· see what you are supposed to be looking for
written? oral?
ask for conclusion, data from - dramatic event?
strong in overt/implied statement
time element
pre-set framework

described. It might also ask her to describe the speaker's attitude, and give evidence of how she determined it. From this written account, the instructor could then assess each student's ability to observe the explicit elements in the poem, and to distinguish what Frost states from what he implies.

To extend the learning potential in this assessment, the second part might involve the student for the rest of the session in a small group, comparing notes to find the basis for each of her descriptions and for her conclusion about the speaker's attitude, and rating them as direct observation or inference. As she works through her notes with her classmates, she will probably notice elements she has not identified which

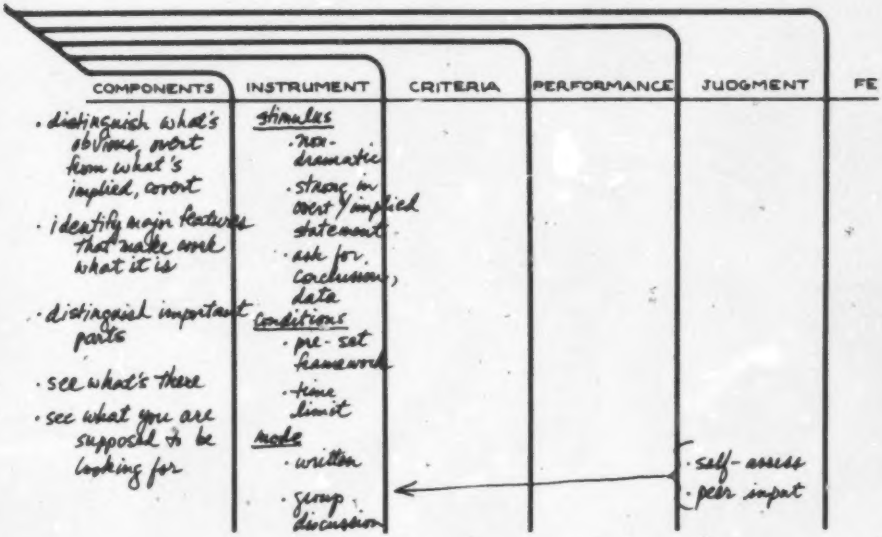
In Hardwood Groves

The same leaves over and over again!
They fall from giving shade above,
To make one texture of faded brown
And fit the earth like a leather glove.

Before the leaves can mount again
To fill the trees with another shade,
They must go down past things coming up.
They must go down into the dark decayed.

They must be pierced by flowers and put
Beneath the feet of dancing flowers.
However it is in some other world
I know that this is the way in ours.

Edward Connery Lathem, ed., The Poetry of Robert Frost; Holt, Rinehart & Winston, 1969.

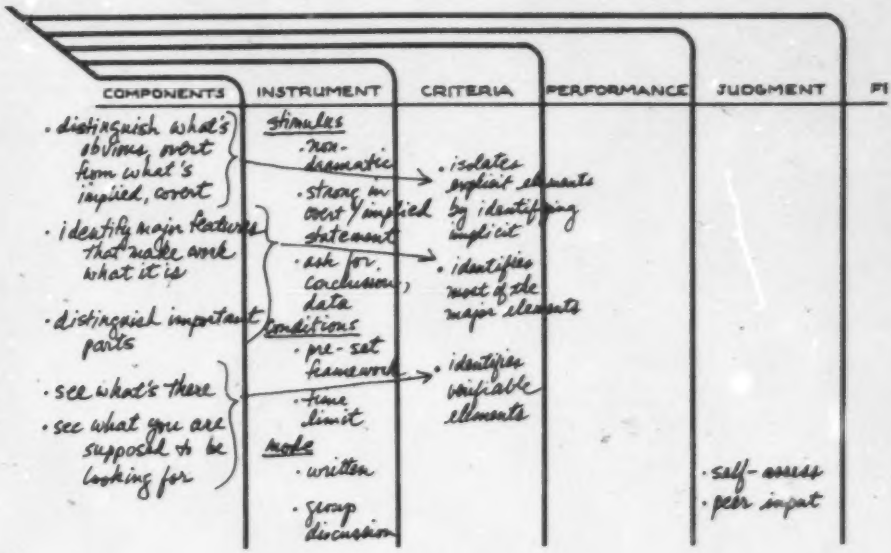


someone else has, or vice versa. Applying the "overt/covert" distinction to her own answers, she can also begin to see where she was using direct observation and where she was using inference. Her notes from this discussion would provide the instructor with a written record of the student's ability to assess her own observing.

The instructor would thus have pre-set the time frames and the framework for observation (things, events, attitude), as well as the mode of reporting or making the "observing" visible (written sentences). This assessment design also engages the student in some self-assessment, with the help of a group of peers.

By what criteria will the instructor judge whether and how effectively the student has observed? This question looks imposing in the abstract, but becomes manageable now that a concrete assessment situation has been defined. How, the instructor can ask, would each of the components I am focusing on be expected to surface in the written record?

If the student is indeed able to "see what's there," the most important component for this assessment, the elements she identifies will include most or all of those that are



verifiable from the text of the poem. Because both the poem and the instrument have led her to make inferences, she will surely have identified some elements by inference rather than by direct observation. Her ability to "distinguish what's obvious, overt, from what's implied, covert" will be reflected in how well she isolates the directly observable elements by identifying those that she inferred, in her annotations during the group work. Finally, her ability to "identify the major features that make the work what it is," to "distinguish the important parts," can be judged by checking to see that her description includes the major elements, however many others she may also have identified.

Having administered the assessment, the instructor would record her judgments according to the criteria and arrange for some form of feedback to the student — written remarks, perhaps accompanied by an in-class review of the assessment for the whole group or by individual conferences, depending on how much of the course she wants to devote to this particular learning experience. Evaluation and redesign would occur in a number of ways. The instructor herself would weigh how well the instrument elicited "observing behavior" and challenged students to develop their awareness and abilities as observers, with an eye to designing further assessments or preparing the course anew. Sharing her assessment design with her colleagues, both informally and through the Analysis Competence Division, she would also receive their comments and make her design available as a model for others.

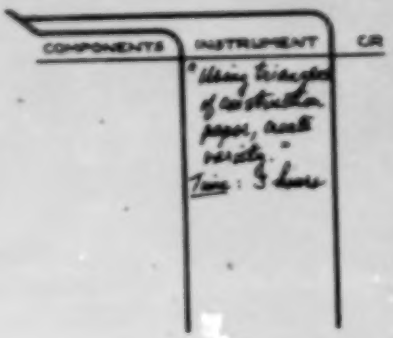
This is how the heuristic model of the assessment design process works to help an instructor move from the outcomes through the rest of the design. As any educator knows, however, good teachers have often been able to "hit upon" effective ways of encouraging students to show what they can do with what they have learned. This is where the model's real usefulness comes in. It is not a fixed sequence, from *A* to *B* to *C*; it is more of an open flow chart that can be entered at any point. Beginning with whatever s/he has most clearly in mind — a good stimulus, some criteria, a particular ability for students to demonstrate or just an unresolved question — an instructor can use the model to help think through and design the other elements of the assessment.

Using the Heuristic with an Existing Technique

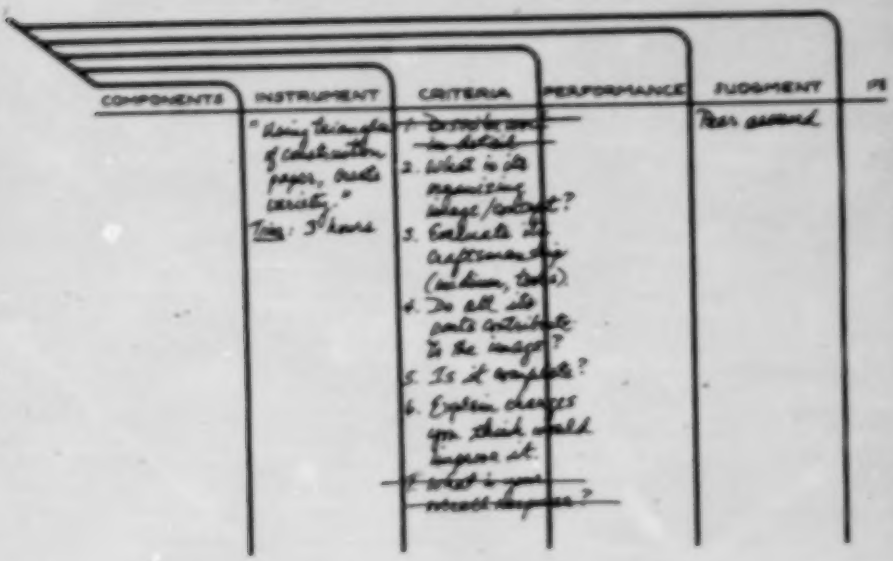
Beginning *in media res* like this, a member of our art department recently opened up some exciting insights into the teaching/learning values latent in the traditional practice of critiquing student work in studio courses. For a class in beginning design, she had readily offered students the opportunity to demonstrate competence in the early levels of problem solving, since such courses commonly proceed by giving students a series of increasingly complex design problems to solve. As she began to work with her opening assignment, however, she sensed an apparent conflict.

The problem she planned to set the class — to create variety, using a single shape (triangles) in a single medium (construction paper) — was a classic exercise in the discipline. But was she pre-empting the student's opportunity to discern and formulate problems by setting the problem in such specific terms and leaving the students only to solve it? Hoping to work this through, she and a colleague went through the heuristic model together to analyze what was really involved in this beginning design.

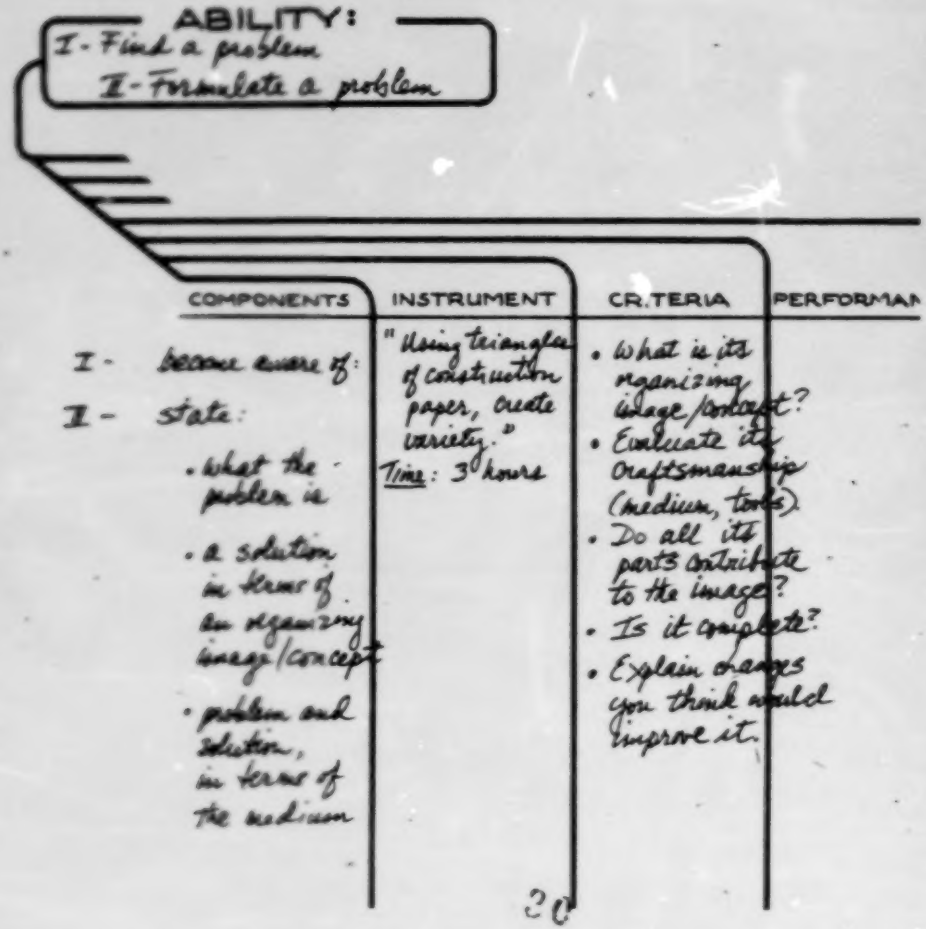
The instrument was already clear.



The art instructor had also been working on a list of steps to guide the in-class critiques, which would be done first by other students and then by her. So the criteria for this assessment might be drawn from that guide list. Indeed, it turned out that her attempts to sketch out what was involved in good critiquing were almost a model "assessor's guide." Setting aside the first and last steps (which are important tools in assessing, but are not specific criteria) each item pointed to a particular dimension of the student's performance in concretely observable terms. Though judgment was often called for, it always had to be backed by pointing to particular elements and details in the student's work.



So far, so good. She then took the salient abilities for both levels 1 and 2 and broke them out into components, to see whether either list of components would "fit" — that is, whether they would be abstract statements of what her criteria were measuring in more specific terms. After some brainstorming and editing, she concluded that "formulate a problem" could be broken down into three major components. These were to state what the problem is, to conceive of a solution in terms of an organizing image or concept, and to state the problem and the solution in terms of the medium. She and her colleague agreed that the earlier ability, "find a problem," might be a matter of becoming aware of these three dimensions prior to being able to formulate them clearly and completely.


ABILITY:
I - Find a problem
II - Formulate a problem
COMPONENTS
INSTRUMENT
CRITERIA
PERFORMAN
I - become aware of:
II - state:
• what the problem is
• a solution in terms of an organizing image/concept
• problem and solution, in terms of the medium
"Using triangles of construction paper, create variety."
Time: 3 hours
• What is its organizing image/concept?
• Evaluate its craftsmanship (medium, tools)
• Do all its parts contribute to the image?
• Is it complete?
• Explain changes you think would improve it.
20

Here, however, was the rub. The conflict she had sensed at the outset now lay before her in detail. It seemed that she was doing the bulk of the students' work for them by stating both the problem and the medium, and was leaving them only the second component — conceiving a solution, an organizing image or idea — to work on. Moreover, it appeared from the criteria that the student was going to be held accountable not simply for conceiving the organizing image but for executing it successfully. This was a more advanced level than she had planned to offer in this course at all — much less in her opening exercise! But even as her colleague summarized this rather bleak picture, the art instructor broke through the dilemma. "That's not true," she insisted. "I don't care at all whether they actually solve the problem."

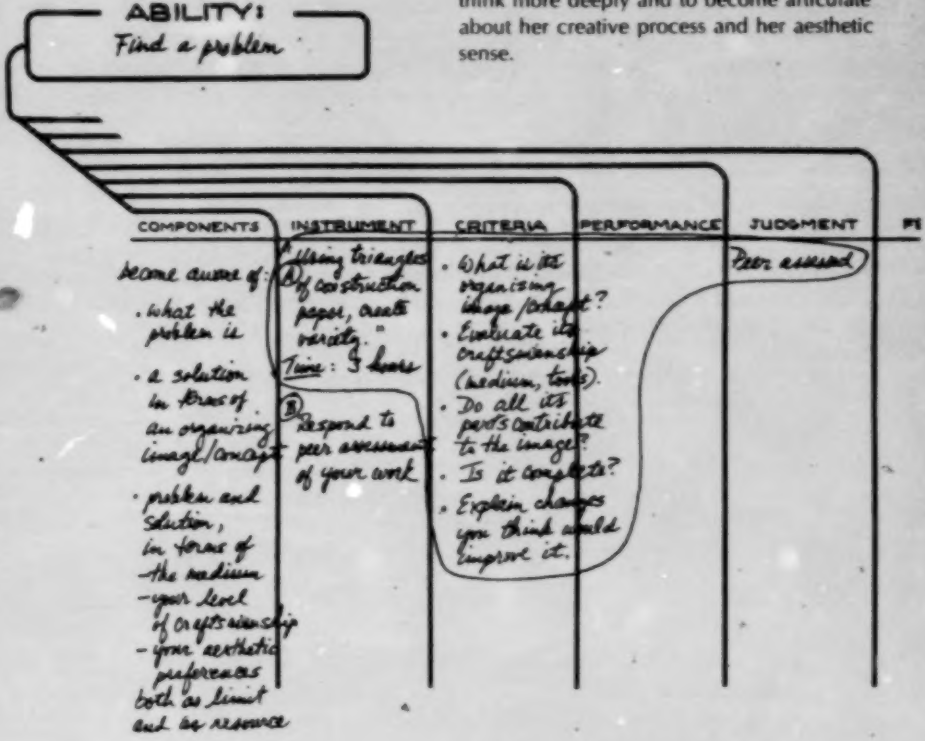
Then why was that among the criteria? To get the student to *articulate* her governing concept or organizing image, to get her to *look at* the limitations of her medium and her own craftsmanship. For many students, the art instructor pointed out, this is a new way of working. They tend to design an ideal solution, unrestrained by the actual limitations of the material and tools before them. Many also are not even aware of having an image or concept in mind as they work. Having to explain to a fellow student what they were "driving at," what vision or feeling was guiding them, reveals to some that they in fact had none, and explains to them why their work has no coherence or sustained impact. For others, it begins introducing them to a sense of their own style, their aesthetic tendencies. Do they respond by drawing out or imposing order, or do they work toward more dynamic states of imbalance? This, she explained, is really the only reason for demanding that the student articulate so much about her intent. It is hardly an essential ability for a practicing artist, but it does help the beginning artist develop her sense of her own vision.

Working these insights back into the model, it became clear at once that the peer assessment was actually part of the instrument. What was being sought in this assessment situation was not the student's solution to the problem set by her instructor, but her responses to the critique of her work by her fellow students. In what she could say and see about whether or not she had worked within the medium, about her craftsmanship in using her tools, about her governing image and her aesthetic tendencies, lay all the "outcomes" the instructor was hoping to foster with this assignment.

The instructor was now able to return to her list of component abilities and clarify it. She broadened the third component to include three kinds of limitation — that medium itself, the student's current level of craftsmanship, and her developing aesthetic preferences — as well as suggesting that the student learn to see these not only as limits but also as resources.

She then began to work on specifying the criteria she would use in assessing how well the student's responses to critiquing would indicate a growing ability to perceive and articulate these issues.

This art instructor was thus able to use the heuristic model to resolve her sense of conflict about her assignment. More than that, it enabled her to probe the many values she had intuitively felt were possible in using the exercise for her class. It also offered the whole faculty a model for peer assessment, or critiquing, as a way to stimulate a student to think more deeply and to become articulate about her creative process and her aesthetic sense.


ABILITY:
Find a problem
COMPONENTS
INSTRUMENT
CRITERIA
PERFORMANCE
JUDGMENT
Become aware of:
• what the problem is
• a solution in terms of an organizing image/concept
• problem and solution, in terms of the medium
-your level of craftsmanship
-your aesthetic preferences
both as limit and as resource
A) Using triangles B) of construction paper, create variety.
Time: 3 hours
B) Respond to peer assessment of your work
• What is the organizing image/concept?
• Evaluate its craftsmanship (medium, tools).
• Do all its parts contribute to the image?
• Is it complete?
• Explain changes you think would improve it.
Peer assessment

Evaluating Assessment Techniques

The instructor's work with the model also raised a question that might well have been passed along to the Problem Solving Competence Division. Is the first level — "find a problem" — a preliminary "awareness" phase of the abilities she and her colleague identified at the second level? Or is it a separate ability to diagnose the tensions in a situation that betray the presence of an unformulated problem? This kind of question, which faculty members frequently encounter as they work with assessment in their courses, is one of the major kinds of information (along with their actual assessment designs) that instructors are constantly feeding back to the respective Competence Divisions. In those divisions, reviewing all these instruments and the questions they raise, the faculty carries out its corporate work of evaluating and redesigning assessment.

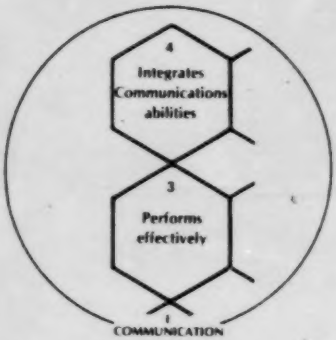
The faculty members in the Communications Competence Division, for example, consider a variety of assessment modes in their semesterly reviews. For level 1 ("assesses her own strengths and weaknesses as a communicator"), the student performs writing, speaking, reading, listening and graphing at the Assessment Center. The "awareness" focus of this level does not demand prior instruction in communications (though the student does receive instruction in self-assessment), and the battery of instruments provides a useful early diagnostic profile of communications skills and needs. This assessment is therefore administered before classes even begin.

In its review of this wholly external, generic assessment, the Communications Division collects overall data on student performance and correlates that with data on the students who took the assessment. Are any instruments or parts of instruments causing frequent difficulty? Are there any patterns of difference between the responses of 18 year old students and returning older women? These are the kinds of questions division members try to probe, with the help of the Assessment Center and the Office of Evaluation's statistical resources.

Division members also check with faculty who offer lower levels of communications in their courses, to see whether students are being placed well and their needs are being diagnosed accurately by what these assessments identify. This in turn helps these instructors to refine and revise their own course designs, as they get a clearer picture of the range and patterns of student abilities and needs.

A different use of the generic approach appears at level 4 ("effectively communicates complex content from a variety of disciplines"). Here, the student must receive one of her three required validations on a college-wide instrument, while the other two come from assessments that are completely designed in individual courses. Drawing the stimulus from one of the student's current courses, the generic assessment for level 4 (done in the Assessment Center) gives her the opportunity to use multiple media to support her presentation.

For level 3 ("performs effectively as a communicator"), the division also has developed a generic instrument. For writing and speaking, however, that means only that they are agreeing on a set of criteria. Each instructor is urged to add further criteria relevant to the particular course and must design the actual stimulus as well. The instructor then briefs a member of the Communications faculty on the modified instrument, and they work as co-assessors of the student's writing or speaking. Because the instruments vary, overall studies of performance data are harder to create than for level 1. Instead, the more common redesign process is used — dialog between the members of the Competence Division and the faculty who offer this level.



So far, the generic criteria for level 3 are still in embryonic form. A single major criterion is applied in addition to those begun at levels 1 and 2: "Treats the topic in a manner that involves analysis and/or synthesis, so that it is more than summary or observation." After a year's work as co-assessors in several course settings, the Communications Division members will meet with the instructors offering levels 2 and 3. They will review their experience and sift the particular criteria added for the respective courses, and arrive at refined criteria for both levels.

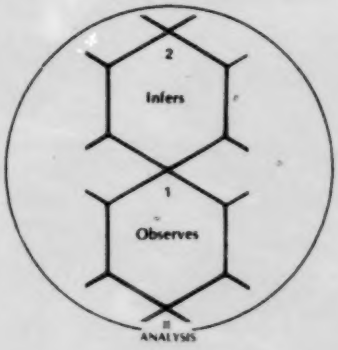
The Communications Division thus uses instruments that are generic in a variety of senses. They moved more rapidly than other competence divisions toward this generic approach, as the faculty discovered that they wanted strong guidance and assistance in teaching and assessing communications skills. Reading and writing, speaking and listening had gradually fallen to English and Speech departments as specialized subjects in colleges and universities across the nation, and had become unfamiliar to the average biology or music theory instructor. By the same token, however, English and Speech instructors who had spent several years wrestling with these abilities and how to teach them were willing to take on the role of "experts" in order to get the creative dialog started.

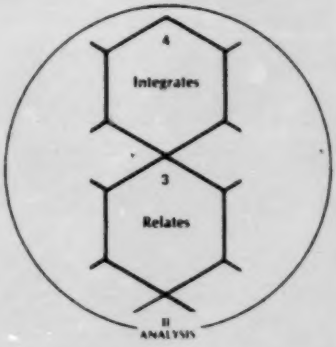
No similar process of attrition had removed analysis, or critical thinking, from its many natural environments. The faculty members who volunteered for the Analysis Competence Division thus found themselves with quite a different task and approached it quite differently. Initially, they offered a tentative outline of six sequential levels for the development and demonstration of analytic thinking. After some discouraging attempts to develop generic criteria and assessment situations, they realized that the "expertise" was diffused throughout the faculty. No one group (themselves included) had spent the years of focused effort on analytic thinking that English and Speech instructors had poured into Freshman Composition and Introductory Speech.

The Analysis Competence Division members therefore turned their efforts toward encouraging each instructor to sharpen her/his expertise. They asked each colleague to work out an explicit model of the kind of analytic thought employed in her/his discipline, and to develop specific assessment techniques and criteria to measure a student's progress in developing it.

Members of the division collaborated individually with 25 colleagues, representing each of the college's academic and professional departments, to capture the current state of the art in the respective fields. They then took a Friday afternoon faculty/staff workshop to report on their findings, and asked two instructors to share their work on how analysis occurs and can be taught and assessed in philosophy and in accounting.

Only as these disciplinary embodiments of analysis become clearer will the Analysis Competence Division begin seeking the generic, college-wide patterns that underlie them. In the meanwhile, they use their semesterly review function to serve as a clearinghouse, linking instructors who are making similar discoveries or tackling similar problems. They also work to broaden the whole faculty's awareness of the many ways in which analytic thought actually takes shape and can be taught and assessed.





Administering Assessment Techniques

An actual assessment instrument may, as the art instructor's did even when she was done, appear little changed from what students had formerly been asked to do on a familiar and proven exercise. It may, as in the case of the poetry assessment, add new elements like the small-group review of the students' written responses. Or it may, like the upper-level integrated assessments for nursing and literature students, involve days of effort on the student's part in activities of a complex and innovative character. What distinguishes even the simplest, most familiar assessment, however, is the presence of such dimensions as the specific component abilities being sought and the criteria according to which the student's performance will be judged. Equally important is the fact that the student's learning experiences have been planned to foster those same abilities, and that she has known from the outset what methods and criteria would be used in the assessment and why.

Whether it is simple and familiar or complex and innovative, whether it takes part of a classroom hour or most of three days in a TV-radio studio or on a nursing home ward, each assessment also requires the participation of one or more trained assessors. It is the assessor who must carefully observe and make a record of the student's performance, noting behavior that reflects the presence or absence of the particular criteria and recording illustrative examples. It is the assessor who must judge whether the criteria were adequately demonstrated to validate the student as capable. And it is the assessor who must either give the student her individual feedback or prepare a detailed analysis from which others may do so.

At one time or another in the student's progress, she encounters assessors drawn from every segment of the college as well as from the wider community. Most often, her instructor assesses her performance, sometimes in conjunction with another member of the faculty. At times, her classmates or advanced students in her area of concentration may also assess her work. She is also always being trained to assess herself, and from the very beginning of her career at Alverno she has met community members who serve as volunteer assessors to provide outside perspectives on her performance.

The several Competence Divisions train faculty and staff members and advanced students who have volunteered to serve as external assessors, as well as alumnae and other professionals from the community outside the college. Both on and off campus, these assessors are recruited by the Assessment Center, which also provides logistical support for their training. New faculty receive assessor training (as well as training in designing assessments) during their orientation. In addition, field and clinical co-assessors are recruited and trained by the respective professional departments, and on-site mentors for off-campus and on-campus experiential learning (OCEL and ONCEL) projects often volunteer to be trained as co-assessors of a student's work.

In each case, assessor training involves a careful discussion of the philosophy and purpose of assessment, and of the criteria fo the particular assessment for which the assessor is being trained. The training sessions also include a number of simulations, in which assessors-to-be practice and refine their observing and recording skills using sample student performances and develop their ability to provide specific, supportive feedback.

Beginning assessors, once trained, are usually teamed with experienced assessors for their first "live" experiences. Assessors are also given "brush-up" sessions at regular intervals, and may be required to engage in such a session if their assessing work begins to fall out of phase with that of their colleagues. In addition, whenever the criteria or other elements of an assessment design are significantly altered, all assessors trained for that assessment are updated with special training.

IV. How Students Experience Assessment

For the student, assessment is one of the most distinctive and powerful parts of learning at Alverno. Assessment is probably her most intense, continuing contact with the Alverno learning system, and it makes education at Alverno visibly different from almost any other college education she might have chosen. Assessment also draws her into an encounter with herself. Any testing situation is likely to be a critically personal experience, a time of maximum vulnerability when a person presents herself and what she knows for external examination. But because the Alverno student's assessments focus on her own developing abilities — rather than on some external norm or objective body of data — they have an even more intensely personal impact on how she sees herself.

Working together as a faculty to evolve a coherent, college-wide assessment system, we have become increasingly able to shape the student's four years of assessments into a coordinated developmental experience. Within the overall system, students take a multiplicity of individual paths and vary widely in their rates of growth. Yet there is a general pattern of development, beginning with the incoming student's first overwhelming exposure to assessment and culminating in the graduate's facility at managing her own learning.

We know the pattern has begun when the student starts **internalizing** the learning process, realizing that it is not "outside" her and understanding what she is doing and why. Her development intensifies later, as she begins **integrating** her several competences into a unified profile of her abilities and her approach to learning and **generalizing** abilities she has developed in one context into other realms of application. By her final semesters, she operates as a **self-directing** learner, planning appropriate learning experiences and helping to design techniques for assessing her performance.

The process begins early — well before the student's first day of classes. Immediately after she is admitted, she selects one of several New Student Assessment Days offered in the months before her first semester begins. The daylong experience, administered by the Assessment Center, begins with a general session setting forth such concepts as competence, assessment and validation. This session reinforces the exposure to these ideas she has had through admissions materials and advising sessions and possibly a campus visit.

Her most effective learning, however, comes from experience. The student spends most of this busy pre-entry Assessment Day participating in half a dozen assessments at the first level of communications. These have a dual purpose: (1) to provide a diagnostic profile of her abilities and needs in writing, reading, speaking, listening, graphing and reading graphs, and (2) to give her an extensive introduction to self-assessment. Although her performance in each of these six areas is carefully recorded for feedback and for use in planning her learning experiences, her validation at this initial level depends solely upon *her ability to assess* her own performances accurately, no matter how successful (or unsuccessful) they are.

For writing and speaking, the student is asked to write a letter and to draft and deliver a three-minute videotaped speech, in each of which she must take a position on an issue close to her experience (e.g., open vs. selective college admissions). For graphing, she is given several items of mathematical information to organize into a single graph, which she then reviews and assesses.

The assessment techniques for listening, reading and reading graphs are all quite similar in design. In each, the student is exposed to an appropriate stimulus — a videotaped speech, an article and a graph, respectively — and is given a series of open-ended questions about it. She then assesses her answers, looking for her strengths and weaknesses.

This beginning communications assessment carries over into her first week on campus. During orientation week, the student spends an hour reviewing the letter she wrote and makes an appointment to self-assess the videotape of her speech. After she has had instruction and practice in self-assessment, she assesses her speech. She then receives complete feedback on her performances. She and an assessor also review her work as a self-assessor on all six instruments.

Students vary in their responses to this first encounter, but they all testify to its impact and complexity. "They had a get-together for all incoming freshmen," one recalls, "to orient me and my family to the new vocabulary, the new system, the whole new outlook. It still left me confused, but it made me start looking. They talked about assessments, they talked about levels and validations and it was just like, 'I'm stepping into a whole new world.'" Others are less eager to begin exploring unknown worlds:

We sat in the auditorium our first day here and the president was telling us, "You're a woman. You're very independent. You're going to learn how to learn in a different way." And everyone's going, "What are they talking about?" It was really threatening and challenging. It was very, very threatening. It was like going into a different culture.

Even with a prior visit to the campus, one junior noted, "It was still really confusing. You have to experience it to really understand the value of it."

Experiencing half a dozen assessments in one day does promote understanding, but it is a bracing way to learn. "You came in for one day in the spring and were bombarded with mimeographed sheets and tests," an art student recalls. "That basically scared me. I didn't know if this is what college was really like, and I didn't know if I wanted to stick to it. But it got better." Nearly every new student would agree with the nursing student who found her speaking assessment the most memorable part of the day: "I never had to make a speech in my life — and we had to give one in front of a video machine! It was the most terrifying experience I had ever had in my whole life."

Counteracting such feelings of confusion and stage fright, however, are the opportunities for self-assessment and feedback. For most students these open up new opportunities, new ways of looking at their own work backed by tangible evidence. "You sit there and watch part of your speech" on videotape, the once-terrified nursing student says, "and you think 'Oh my God, I did this and this and this wrong,' and you don't concentrate on what you did well. But they stress that more than they do your weaknesses. I think that's really important, especially when you're just beginning." A former high school debater found herself "very excited to see that you can break it down into all these behaviors and parts of the speech, which I had never done before. I thought by 'evaluation' they would tell me 'your speech was good,' or 'Your speech was bad.' But they really made some concrete statements about it."

Getting Used to Assessment

After classes begin, the student's experience of assessment broadens rapidly. In her first semester, she probably will at least choose to attempt validation at the first level in all seven remaining competences.

Three of these assessments will occur in the Assessment Center, in conjunction with the required Freshman Seminar. For her beginning assessment in social interaction, she joins five other students in a task-oriented group simulation. While a team of six assessors observes and records each student's verbal and non-verbal behavior, she and her five peers work to reach a consensus. Afterwards, she outlines the 30 to 40 minute interaction as best she can from memory, and evaluates the group's achievement of its goals. She then assesses her strengths and weaknesses, and those of each of her peers, according to specified criteria. After the external assessors reach consensus, she meets for a one-to-one feedback session with a member of the assessor team.

By contrast to this team experience, the student takes her first assessment alone in the environment competence. Using a written guide and a videotape, she identifies several types of environmental systems (natural, sociological, economic, aesthetic, etc.) within various settings and selects one for a more detailed descriptive analysis. Her written transcript is reviewed and annotated by a trained assessor, with whom she has the opportunity to meet for individual feedback.

The student works with two peers in her first contemporary world assessment. Within a pre-set global issue, such as energy, she and her two teammates select a more specific topic (solar energy, opening up offshore US reserves, etc.) and each takes a separate "interest focus" (oil companies, consumers, government and so on). After several weeks of research the students appear as a panel and are interviewed by a trained discussion leader, while a team of three assessors records their remarks. Each student has a supplementary one-to-one followup session with an assessor. The three assessors then reach consensus about each student's awareness and involvement in relation to the topic, its implications and its importance for her, her country and the world. She receives written feedback, with the opportunity for a personal session with the discussion leader.

For each of these three assessments, the student receives extensive preparation in the Freshman Seminar. Class sessions acquaint her with the concepts and terms involved — the task-oriented model, for instance, and the various kinds of behavior (initiating, summarizing, challenging) that can be isolated in a group discussion. She and her classmates also receive repeated opportunities in class to learn them in practice. The student then decides when she is ready to attempt validation, and makes an appointment with the Assessment Center staff. Should she not be validated, she may make a new appointment and re-take the assessment until she succeeds in demonstrating the abilities involved.

For the first level of each of the remaining competences, she is assessed in her various courses. Because assessing in multiple settings is required, she may be taking as many as a dozen or more separate assessments in order to be fully validated at level one. In a beginning chemistry course, for example, she may be offered learning experiences and assessments for the first level of analysis and problem solving. In a music history class, she might seek another validation for analysis (for which a total of three is required) and one for aesthetic response (for which three are needed).

In her courses, as in the Freshman Seminar, numerous repeated learning experiences help the student develop her understanding and ability to the point where she is ready to attempt validation. These learning experiences frequently incorporate elements of assessment, such as specified criteria, feedback, and so on. Many instructors also use an assessment at the beginning of a course, to establish a kind of baseline for students to use in judging their own progress.

Once she has been validated at the first level in a given competence, the student may also work toward level two of that competence — in the same course, if it is offered there. Indeed, many beginning freshmen contract for a dozen or more second-level validation attempts in various competences. A student may thus participate in more than two dozen assessments for validation in her first semester — not to mention the many diagnostic or formative assessments her instructors may give. She also will be called on to assess her instructor's performance in each course, as part of the regular faculty evaluation process.

As imposing as this sounds, she actually spends no more of her time involved in assessment than her predecessors used to spend writing papers and reports, "cramming" and taking exams. Unlike traditional tests, however, each of the student's assessments contributes directly to her learning. Within each course, the assessments are designed primarily as learning experiences, helping her to become aware of and to develop abilities fundamental to the discipline. In addition the overall breadth and number of her assessment experiences, both in class and in the Assessment Center, provide her very early in her college experience with a wide-spectrum view of her own abilities and needs.

The Internalizing Phase

As a result, these first few months enable the student to begin **internalizing** her learning experiences. She moves away from things that "they" expect her to do in order to "get by," and toward the understanding that all this is happening to her and for her — that she, not the instructor or the material or the requirements, is at the center of the process.

The rapid total immersion of her first few months actually accelerates the student's growth toward a personal understanding of the Alverno learning process and herself as a learner. She is so extensively involved, in so many diverse areas and with so many parts of her own self-image, that she assimilates the entire process and makes it her own much more readily than she would from a more restricted, gradual exposure.

When our assessment system was in its infancy and offered fewer, less complex alternatives, students regularly reported an "Aha!" experience — suddenly comprehending what they were doing and why — in the second or third semester. As our system has matured and the range of options has multiplied at every level, most freshmen are reporting that breakthrough in the first month or two.

"At first glance," one student recalls, "it's frightening. You came in wondering exactly what in the world a competence level was and you never really understood it until halfway through the semester it dawned on you like the sun coming up in the morning. It just hit you all of a sudden — 'Oh, this is it!' — and it seemed like something you should have realized a long time ago." As another student put it, "The competence system isn't something that can be taught. I noticed that in my freshman year, when I had to do a philosophy assessment. It's something that can't be taught but just hits you."

The breakthrough isn't always an unexpected gift; many work hard to achieve it. "A few weeks into the semester," one sophomore says, "I realized I didn't know where I was going. It wasn't coming together. So I just sat myself down with all the syllabi and the materials they hand out and I read. I read for a whole day until it started making sense. I haven't had any real problems since then." Others struggle longer:

I remember a whole semester running around trying to find someone to explain it to me . . . somehow or other I guess I sort of figured it out. It's a painful process in the beginning when you see those horrible words, but all of a sudden you get this practice and the process is practice. It's the practice, I guess, that whole thing of newness, then practicing it, and then it becomes a part of you.

As she contends with these demanding new experiences, trying to pull them into some meaningful order, the student can build up a strong sense of frustration. The unusual elements of her education — assessment in particular — offer ready targets. "When I first came here I was pretty unsure about the program," one student reflects. "I guess I did go through a period of negativism where everything about it must be wrong." "I hated what they were telling me," one senior says bluntly. "If you would have asked me as a freshman what I thought of Alverno and the assessments, I probably would have slugged you."

Because ours is a college-wide system — designed and evaluated by the whole faculty working cooperatively — we are much better able to tolerate and even to plan for such pressure points. Frustrations and pressures are, after all, inevitably part of any task worth undertaking, and they are especially critical to personal growth. A student used to passive and abstract learning may well feel unfairly put upon when she must take the initiative and actually engage in an active learning experience or perform in an assessment, instead of simply reading a text or playing back lecture notes. Similarly, facing up to her performances and what they say about her abilities and weaknesses is likely to be difficult, threatening work for a student whose prior schooling has led her not to be reflective at all (as in rote learning) or to be irresponsibly subjective.

In a traditional setting, no single teacher or department could get away with demanding so much, when others offer easier alternatives. But our joint effort to create a unified atmosphere, in which the student "bumps into what she needs no matter where she turns," allows us to greet signs of appropriate frustration as indications that the growth process is moving along well. As one department coordinator put it, "I know we're on the right track when the freshmen come to me and complain, 'Why is that instructor getting paid? *I'm* doing all the work!' "

Along with built-in frustration, of course, must come built-in support. For the new student this takes several forms. The Freshman Seminar, for example, provides a forum for ventilating problems and finding ways to cope, as well as putting each new student in touch with a peer counselor who has weathered the storms she is going through. The students also give each other important informal support:

There were times when you just couldn't do some of the things they would expect you to do. Or else you would do it and work your heart out and your soul out and they would say,"Well, that's not right. You missed the whole thing." And you go,"Well, what am I supposed to be doing?"

We'd have group sessions, and we'd look at these competence levels and try to figure them out together. We all worked together and we leaned heavily on our instructors, who had to lean on us also. Of course they wanted to see what we were doing as anxiously as we wanted to get it done. We really had a hard time.

In addition, the faculty pays close attention to student expressions of discomfort and stress, although they begin to sound familiar and even welcome at the right times. They are, after all, among our surest indicators when pressure exceeds productive levels or occurs out of phase, requiring us to look immediately and carefully for ways to modify what we are doing.

Assessment itself, however, has proven the most reliable antidote for the anxieties that assessment arouses. At first, the student may be overwhelmed by re-reading an awkward essay or watching an embarrassing case of nerves on videotape; positive feedback may fall on temporarily deaf ears. But repeated assessments soon help her relax and grow accustomed to scrutinizing her work critically. "If you are lacking in a specific area, they'll tell you," one student explains. "They'll work with you to improve that area, but it doesn't mean you pass or fail. That takes a lot of stress off the kids here — the idea that if you don't exactly do it right the first time you try again, and you're given specific things on which you can improve."

Gradually, these "areas" move closer and closer into the student's perception of herself, as she literally internalizes their meaning. This eventually provides the most powerful positive feedback of all: "I got to a point where I realized I was learning from it. After the first couple of months, I really felt that it was doing me good." "Once you get into the system," a sophomore recalls, "you realize, 'Well, now I can do *this* and do *this* and do *this* — and I couldn't when I first came here.'"

This discovery — that what she learns isn't poured into her, but is being elicited and cultivated from her own innate capacities to understand and to do — often sets in motion a thorough realignment of the student's self concept. She gains a new way of looking at learning and also at herself. "It takes you that whole first semester," one sophomore says, "to get acclimated to concentrating more on your skills rather than just on content. You're just getting in touch on your values . . . on the idea that you do know something about humanities, and you do know something about environment."

"It's not something you're going into blindly," another agrees, "like you think it is when you first see that matrix of competences and wonder, 'How in the world do I get these skills?'" A senior's poignant reflections sum up the experience of internalizing:

When I came here I thought that everyone had the knowledge and I had to find the people who had the knowledge in order to get brains, education, talent, whatever. As I have gone through the Alverno learning process I've struggled through some things, and I've realized that the stuff is inside me and that people can help me bring it out and hone it down so its rough edges are gone and considerably expand it.

I think that's what Alverno is doing. They pull it out of you. I mean you're kicking and yelling and screaming all the way, but somebody keeps pulling that stuff out of you. It shocked me to find I have a brain. I'd never thought it in a million years. And that's what's changed about me, that I thought everyone had it and I was the dummy. You get that self-confidence thing going. It's the greatest to say, "I can do it," or "I think I can do it," or "If I can't do it I'm going to know why."

The Integrating Phase

By the end of her first year, the student has usually been validated at the first and second levels in nearly all the competences. She spends most of her sophomore year working at levels three and four, and may do some fifth level work in competences related to her area of concentration. Again, because most of these levels require assessment in multiple settings, she may earn as many as three dozen or more validations during the year. By the third and fourth levels, however, we have begun combining assessments: on the initiative of either the faculty or the student, a single assessment may be used to validate for two or three competences at once. This practice reduces the total number of her separate assessments while increasing the complexity of each of them.

A recent course on the comparative history of revolutions, for example, offered learning experiences and assessment opportunities for four separate competence levels — contemporary world, levels 3 and 4, and aesthetic response, levels 3 and 4. The course began with lectures and discussions providing an in-class analysis of four historic revolutions, using an analytic framework derived from Thomas Greene's Comparative Revolutionary Movements. The class also looked at literary and graphic arts from the four societies under study, to see how revolutionary change might influence the ways individuals derive and express their identity as persons. Students then worked in teams to identify and study a current revolutionary situation, dividing their tasks according to the competence levels each team member chose to work on.

Several formative assessments were given during the first portion of the course. Each student's final assessment included participating in an hour-long panel presentation by her team and completing a worksheet packet for each of the levels in which she was seeking validation. Thus, one student on a team studying South Africa made her oral presentation as part of her assessment for level 4 in aesthetic response. First, her teammates on the panel had described the South African racial situation and its causes (level 3 of contemporary world), compared the post-World War II racial developments in the U.S. and South Africa (level 4) and read and discussed a poem by the South African black poet Mongane Serote (aesthetic response, level 3). She then took Serote's poem and compared it to another, by the American Langston Hughes, pointing out the common elements and analyzing the differences in terms of the two authors' respective cultural and political milieux. She also submitted written worksheets for this level, as well as for level 4 of the contemporary world competence and level 3 of aesthetic response. From this single assessment experience, she received validations in all three competence levels.

It is also not uncommon for the student herself to suggest ways of combining her assessments — using a teaching project she develops for an education course, for example, as an opportunity to display ability in several media for one of her level 4 communications validations. Instructors often urge sophomore students to seek such "doubling" effects. Even the assessments administered through the Assessment Center are woven more closely into her classroom experience, all but one of them being co-designed and/or co-assessed by an instructor from one of her courses.

The effect of this emphasis on combined assessments is to encourage the student toward **integrating** her learning experiences. As a faculty, we are aware of having separated the several competences and their sequential levels in order to teach and assess them. For the student, keeping these abilities discrete from one another at first helps her to focus on abilities she may well have taken for granted and to see that they can be analyzed and consciously improved. (Typical is the student already quoted, who recalls discovering with excitement that "you can break it down into all these behaviors and parts of the speech," or the one who says of social interaction and valuing, "You know, talking to people and knowing where you stand just seemed to me like something people did automatically — didn't think about; didn't have to think about.") Now, however, combining assessments gives her a direct experience of how the competences interrelate. She begins to realize that while she has learned to identify and work on her developing abilities one by one up until now, they actually function inseparably in her own performance and growth.

This dawning sense of integration is likely to emerge first in her understanding of how the levels in a given competence work together. The cumulative nature of the levels appears graphically in the criteria for assessments, where each successive level includes criteria for the prior levels *in toto* or in summary. As her experience of successive assessments brings the meaning of this repetition to her, she is likely to say things like, "After you've mastered the first level and you go on to the second, you're still doing all the things you did at level one. You can't forget it because it's integrated so closely in the next level, and so on it goes all the way down the line."

She also begins linking competences together, as she sees one ability — say, inferring implicit elements (analysis, level 2) — having a direct impact on another — such as discerning the values implicit in a technological decision (valuing, level 3). "Actually," one student explains, describing a larger network, "social interaction ties in very closely with the communications and the analyzing and the problem solving, because you put all of them together when you communicate." Or as a nursing student describes it, "As for the idea of the competences, everything is totally integrated — in our learning experiences at the hospital, in our tests, everything. When you're analyzing, you do your social interaction . . . it's not what it was before on your lower levels. It's a lot less pressure because everything you do is very intertwined."

After repeated nudges, the student's growth toward integrating her learning receives an outright shove at the end of her second year, when she participates in the Integrated Competence Seminar (ICS). This half-day assessment, administered through the Assessment Center, specifically focuses on her ability to demonstrate in an integrated manner the several competences she has been developing. She and four or five other students are assigned roles within a fictitious group trying to deal with a civic problem, and each receives a different set of background material. From the alternatives proposed, each student must make a choice and then give an eight minute speech advocating her position. Each is then given a separate "In-Basket" instrument, in which she must sort through a bundle of tasks, memos, letters and phone messages that have accumulated in her "office" and decide which ones to deal with and which to delegate. After she has completed that forty minute exercise and a rationale statement explaining her decisions, she and her colleagues meet as a group to decide which solutions they will implement.

This is an intense, compressed experience. The "In-Basket" is specifically designed to include several "forced choice" either/or conflicts and has so many elements that not every item can be responded to within the time frame. The group simulation likewise operates under firm time constraints, and during it each student discovers that she has been given only a partial view of the necessary information for defining and solving the group's problem. The ICS experience also differs sharply from the pattern of assessment she has become used to, in that no specific learning experiences in any of her courses have explicitly prepared her for the problems and tasks she encounters. The Integrated Competence Seminar is thus a wholly external assessment in its relation to her prior learning experiences, and predictably generates a good deal of anxiety and frustration.

At the same time, however, the ICS provides a valuable insight into how ready the student is to pull her abilities into integrated operation on a new kind of problem with the "real world" characteristics of *not* being separated neatly into its elements and *not* coming as the culmination of a series of preparatory learning experiences. It also gives her a dramatic experience of how her several abilities interact. "The sophomore year we had to do this 'Integrated Seminar,'" one senior recalls. "You took a position, and you had to set your priorities because you had a stack full of things on your table. And I thought, 'Wow! You're a woman, and you should appreciate a business person more because of all the things they have to do!' And then I started to pull together the different qualities I needed."

As the growing emphasis on combined assessment culminates in the ICS experience, students find integrating becoming more and more habitual. Indeed, it becomes a dominant mode of their upper division experience. "At first," one senior recalls, "everything seems separate. But when you get up into the upper division, things start to be put together so that when you're giving a speech you're also analyzing the situation, you might be problem solving . . . you're taking everything into consideration. You might not be touching on all eight all of the time, but you're using more than one. They really do fit together." Another senior says, "They break everything apart so that you do get to know yourself, what you might value, how your methods of interacting might be. But then by the time you get to be a junior or a senior everything gets put together and you're putting together your values along with your knowledge so that you're one person, so you can go out into the world and have everything together."

The Generalizing Phase

The Integrated Competence Seminar's "unprepared" externality also sets in motion another pattern that becomes increasingly important during the student's third and fourth years. It spurs her to begin consciously **generalizing** her abilities beyond the situations in which she has learned and developed them. In so doing, she is making explicit what has been an implicit element of her experience all along. Working towards analysis level 4 in a philosophy course, for example, she was drawing on and extending the analytic abilities she had developed at levels 1, 2 and 3 — though they might have been learned and validated in a combination of math, art history and Spanish classes. The ICS experience focuses her attention on this transfer effect by requiring her to function effectively in a situation foreign to any of her course experiences.

As she begins working more closely in her chosen area of concentration, this transfer effect intensifies. Abilities she has developed in a broad variety of settings must now be brought to focus in a more sharply defined professional context. This is particularly true as she participates in clinical or field experiences appropriate to her profession. A junior nursing student preparing care plans for her clients, for example, draws on and develops the abilities she has already shown at levels 1-4 in problem solving, valuing and social interaction. Not only is she transferring these abilities from the several contexts in which she originally learned and demonstrated them, she must also now generalize what she has learned from working with healthy client populations and apply it to the needs of clients suffering from acute illnesses.

For the liberal arts student who is not in a professional program, off-campus experiential learning (OCEL) projects or their on-campus (ONCEL) counterparts provide similar experience in generalizing. "Every week I had to fill out a log," recalls a co-founder and former editor of the college newspaper, who used the experience as an ONCEL project.

At the top it said, "Check which competences you applied here." Well, every week I checked just about every one. And I could give specific examples . . . because I was obviously using my communications skills, obviously I had to analyze and do problem solving, and valuing? There's a lot of valuing involved in setting up anything that has to do with people and institutions. Social interaction? Well, obviously I was interacting with my staff, I was interacting with the printer. . . . Environment? I was working with the social environment of Alverno, and within the physical environment of the places where we did our work. We were trying to deal with some current events, and so on.

It just made it really evident that I was transferring my abilities, having to write down which ones I used.

Predictably, this generalizing pattern extends beyond the student's learning experiences, on or off campus, and she begins to use her developing abilities in every aspect of her life. "You have to look at it so much that it becomes a part of you," one student says. "Kids around school will joke about it, and will say, 'Oh, now she's exhibiting withdrawing behavior.' But when I go home and somebody says, 'Gee, I don't know if I should go here or there' I find myself going through all the problem solving stages with them. Maybe I don't label them, but I know where I'm at in each step."

Not only do the processes she has learned carry over, but the habit of assessing her own performance and that of others becomes generalized as well. "At the end of any interaction," an education student reflects, "you always look back on your own and say, 'Now did I do that right? Could I have been more effective? What could I work on for next time? What did I do extremely well?' It's just become second nature, really." "I'm involved in a group outside of school" another student says, "and I've been going to that committee meeting and I get so disgusted with them because it's been ingrained in my head about meeting formats and how to start and plan your agenda and brainstorm . . . It's just part of me now and I go to these meetings and it seems like total disorganization."

An important payoff for the student — and for the faculty as well — comes when her generalizing of these abilities helps her handle her life with increased effectiveness. "At Christmas," one senior recalls, "someone was in charge of liturgy and wanted to have a sing-along in the commons, around the tree. Well, you had dorm students, you had cash paying customers, you had customers with tickets . . . I said, 'Well, Ann, did you think of this and this and this?' She just looked at me and said 'No. What am I going to do?' And I sat down with her and we effectively solved her problem." One student recalls applying self-assessment during a critical decision:

I was about to drop out of college because I didn't successfully complete some nursing courses. Quit – that was my first impulse. "I didn't succeed in one thing, I'm not going to succeed in anything."

"Well," I said, "that's wrong." And I guess I began critiquing my decision. "How is it going to affect me today and tomorrow?" When you have to make life decisions, it's hard, and naturally you have to critique on how you're doing this far, how it's going to affect you in the future. But I don't think normally I would have done that. I do it more now than I did before I came here.

The Self-Directing Phase

The final phase of the experience, as the student moves through the college-wide assessment system, appears as she becomes an increasingly **self-directing** learner. This is specifically fostered by the very nature of assessment, even before she enters the college, in that she is constantly called upon to assess her own strengths and weaknesses.

A senior in management tells of using her ability to assess her own growth and needs and to interpret an environment in choosing her senior OCEL project. "I had three different places to pick from. I picked the problem I wanted to solve, the one that was going to challenge me the most and let me show my successful management abilities." "I can see it in my family," another student says. "I'm always trying to bring these skills home, to stimulate the kids, even trying to teach them. I also use it in the hospital, I use assessment everywhere. I mean, you can't just use it here and not use it in other situations. Once you've kind of integrated it, then you use it all the time."

As she progresses, she is also invited to take a greater and greater share of the initiative in deciding which learning experiences she needs, and in what sequence and pace. "You sign up for when you're going to do your communications level 3," one student explains, "and you do it in the class that you choose. Most teachers don't badger you to get it done, you have to take the responsibility. But if you run into trouble, there's nobody saying, 'You *can't* go to this teacher or that lab instructor.' You go out and seek her on your own. If you don't get validated, you really have nobody to blame but yourself." "Now I'm a junior," another says, "and I'm going to do my speech for communications, level 4. You do that totally on your own unless you ask for help. You know what your weaknesses are because you're constantly assessing yourself. They're more or less just reinforcing your own assessment skills."

In her upper division work, the student participates increasingly in the design of her learning experiences and assessments. Not only does she choose her own OCEL or ONCEL site, for example, but she is responsible for co-designing the assessment techniques and criteria by which her performance — both her work for her employer and other aspects of her learning — will be assessed. "Senior year and even junior year," one student recalls, "you have to do all these projects and you learn to be really independent and to be organized and to do things on your own. Otherwise they're not going to get done. That's one thing stressed around here." "There are times when we have to plunge into things," another senior adds, "and we don't know what we're doing. There have been a number of upper level competences that I've done that we had one foot on a banana peel as far as guidelines. I mean there's a risk involved."

By the time she is facing graduation, the senior begins to project these self-directing habits into her inevitable musings about her future career. "I think we're always going to be assessing ourselves," one nursing senior says. "I know that before I leave the floor at the end of the day, I'll make myself notes just to make sure I've done everything."

"Assessing has become such a part of you," another reflects, that it's something you're going to do your whole life. Every time you do something, or when you're about to do something, you're going to think about it. And after it's done you'll say, 'Well, what did I do well? What did I do wrong?' You're going to use it everywhere — in your nursing, in your art, in your marriage. It will be integrated into our lives." A third student sums it up by saying, "You're your own private Assessment Center by the time you get out of here. And you can find people you need to help you. You create your own resources."

For the student who has realized the full benefits of the assessment experience during her college years, assessment has become an integral part of her learning just as learning has become an integral part of her life. "Now I know," one senior says, "that there'll never be a time when I'll stop learning." "Somewhere along the way I changed my attitude," another says. "Assessment — learning — what's there to be afraid of? It's all a challenge, your own personal development. It's fun!"

"I'm not so worried about comparing myself to the next person," another senior adds. "I compare myself to myself. As long as I know what to do to learn, that's important." "I used to approach learning with a lot of fear," another says simply. "Now I know I can make the most of it."

V. How We Develop Our Assessment System

Often during the last few years, we have been asked such questions as "How do you set up an assessment system?" or "What is the organizational structure of Alverno's college-wide assessment program?" We have even asked ourselves whether we could describe our model in overall terms, and have tried to do so. The more deeply and seriously we have reflected upon such questions and upon our own experience, however, the more we have become convinced that our most important learning has been to see our "system" as a constantly evolving process rather than as a structure.

When we first began our exploration into assessment, we looked far and wide for anyone who had dealt with the problem of developing assessment techniques. Collaborating with the creators of academic assessment practices, nascent programs in the professions and several well-developed business assessment centers, we discovered that the most unusual aspect of what we were attempting was our intention to use assessment as a major developmental experience in the learning process.

Employers, educators and credentialing agencies valued assessment as a way of looking systematically at "realistic" performance abilities directly related to the job role or profession. In response to the needs of their environments, assessments were often being used much like traditional tests — for identifying, certifying and sorting candidates. The more we and other assessment developers began to use assessment as an integral part of the individual's ongoing learning, however, the more complex the task became.

For one thing, the sheer size of an assessment system designed to follow several hundred learners through four years of their development so far exceeded anything then in existence that it was no more possible to describe in meaningful detail than an unexplored continent. In addition, our commitment to liberal education involved us in assessing a broader range of abilities than most previous assessment ventures had needed to deal with.

Even more important was the fact that in a developmental context the goal is to create a system responsive and flexible enough so that each assessment provides the maximum possible insight into the individual learner's progress. Although intake assessments may be fairly standardized and universal — as is our pre-entry Assessment Day — each student's progress beyond that point will necessarily tend to become more *individualized*. Because the initial assessments reveal an individual's particular strengths and weaknesses and learning style her pattern of courses and learning experiences can be selected to suit her particular needs. Subsequent assessments must, in turn, reflect the individual pattern of learning experiences she has engaged in, and the developmental goals for which those experiences were designed.

In our developmental use of assessment, moreover, we specifically wanted the entire learning process to become more and more *learner-directed*. This goal, too, affected the nature of the assessment system. We had already learned from the designers of business assessment centers that it is critical to inform each candidate beforehand of the purposes, methods and criteria of the assessment and to provide detailed personal feedback afterward. Extending this process to train the learner to select her own learning experiences and assessment situations and to assess herself augmented the need for a responsive, flexible system.

Finally, our developmental emphasis required that we take assessment as seriously as we ask students to, and that we rely on it as a primary source of feedback on our goals and efforts as educators.

Frequently, a pattern of student difficulty with part of an assessment leads instructors to redesign their instruments. From time to time, a pattern of unsuccessful validation attempts reveals a failure to teach as well as a failure to learn, and prompts a rethinking of a learning experience or a whole course. Thus, for example, the Freshman Seminar's learning experiences for level 1 of the environment and contemporary world competences have been thoroughly reworked, along with the assessment situations themselves, in response to student difficulties that emerged in those assessments.

Conversely, successful and creative student performances are among the most gratifying forms of feedback for faculty. Consistently impressive performances on an assessment have on several occasions encouraged instructors or departments to extend the offerings in a given course to include more advanced levels of a given competence or to embrace additional competences.

Using assessment developmentally in undergraduate liberal education, then, means that each of more than a thousand students spends four years pursuing a personalized learning path. Each instructor or department designs and modifies assessments to fit the learning experiences within each course, and the learning experiences also change from semester to semester. Upper division students regularly engage in independent projects or in field experiences shaped by needs and opportunities in the community, for which learning and assessment plans must be adapted or improvised. And throughout her progress, the student plays an increasing role in selecting her own learning experiences and assessing her own performance.

A developmental system of such scope and constant flexibility goes far beyond any of the assessment systems developed before we began our work. Nonetheless, to manage such a system the faculty must be able to assure that its fluidity operates within a unified, coherent framework, and must be able to maintain certain levels of coordination and quality.

where its impact may be important. They also work to maintain a general consistency of development throughout the system so that no one area or element slips out of focus, as in their recent effort to encourage wider use of oral and multi-media assessments throughout the college.

Elements of Change

The Assessment Committee's most important responsibility, however, is to carry on a constant rethinking of our goals and our progress towards them and to seek out and nurture any promising new avenues of inquiry or insight. Just as the Competence Divisions do with individual competences, the committee members work constantly to pull together what we have learned about assessment and its contribution to learning and how we might wish to alter or augment our overall assessment system. In this role, the Assessment Committee functions not only as an agent of synthesis and coherence but also as a source of continuing fluidity and change.

A significant instance of this arose recently as individual departments began to crystallize some patterns they were seeing in student performance on advanced assessments. As students neared graduation, instructors in every area found themselves facing the question of how to make verifiable distinctions among student performances, and between those that were successful and those that were outstanding. What criteria, they asked, were they really using in sensing and agreeing on such distinctions?

In the mathematics and biology departments, for example, instructors had begun to assess for "independence." On assessment records detailing criteria for each step in the scientific problem solving method, biology assessors were also noting whether the step was student-initiated (+1), faculty-initiated (0) or required repeated faculty intervention (−1). The math department had developed a ten-point scale for rating the independence of a student's work in each of five areas (definition, axiom, theorem, computation, application). Other instructors were making similar efforts to become explicit about their apparently intuitive judgments in discriminating among successful performances and describing the personal styles of individual students.

Aware of these and related concerns and their potential implications, the members of the Assessment Committee were working to define the elements that distinguished advanced-level learning from a student's earlier development. At the year-end institute, they shared their work with the faculty and staff on defining six "characteristics of advanced student performance."

What is your dialectic? Whole faculties, even at a small college, are too large to do much of the detail work that must be done in bringing new conceptions into clarity and clarified ideas into reality. Committees and task forces are too small to think of all the relevant possibilities or to generate and maintain faculty-wide consensus. Some sort of ongoing dialectic was the dynamic we found essential for accomplishing all these goals simultaneously. At one point the dialog was president-faculty, then it was largely task force-departments, and now it occurs in numerous modes (Competence Divisions-individual faculty, Assessment Committee-Competence Divisions, any group-whole faculty via workshops and institutes, etc.). The creative dialectic can take a protean variety of forms; but if it isn't there, nothing happens.

Are all your definitions open? We are constantly astonished at how perishable insights prove to be. No matter how carefully we formulate our understanding about one element of the system, within a year or two the things we learn in other areas reshape that summary's meaning so that it looks at best incomplete, at worst positively inaccurate. Not one of the competences or its constituent levels, not one of our statements of committee duties or priorities, not any assessment technique or learning experience doesn't soon need revising or rethinking. Everything can't be rethought at once, of course; and some revisions are more urgent than others. But any time a statement or a set of criteria more than a year old looks as good as it did the day it was written, look again.

What are you trying to assess? Whether we are rethinking some part of the existing system or wrestling with a new problem, we always find ouselves back at "Go." We simply cannot work effectively with any part of the assessment system until we have carefully stated (or restated) which outcome it is supposed to help us assess. Often, when we reach a roadblock in our thinking, we find that it is because we are really as far as the old statement of the outcome will take us, and we are readier than we know to restate it.

Who are your assessment "naturals"? Assessment isn't a field in which anyone has trained expertise, not even measurement psychologists. But your system will demand the whole professional effort of *at least* one or two people who find that they actually enjoy reading, writing, eating and breathing assessment. We have found home-grown talent to be essential here, particularly in the earlier stages. It would be futile to import a system or follow an outside expert's formulations: you can only assess what *you* mean to teach. At some point, your faculty's assessment specialists will find themselves diverging from whatever professional paths they had been following to shepherd the system's growth. But if they are indeed "naturals," they won't mind.

Have you looked at the whole picture recently? The process of thinking through and planning any part of the system — or of successfully designing and carrying out an assessment that reveals genuine learning progress — can be both engrossing and rewarding. We have had to remind ourselves at every level to keep looking at the overall view, to re-evaluate goals and directions, to pay attention to the total "shape" of what we are doing. This is a critical job, and because it doesn't always seem pressing it needs to be constantly reinforced as a priority. But we have found that it pays some of the richest dividends in helping us realize our accomplishments and find promising new avenues of development. It can also help stave off cumulative errors or omissions.

Like anything worth doing, developing an assessment system (or any part of it, even an individual assessment technique) isn't easy. As an ongoing system, assessment makes demands on student and faculty time that are roughly equal to what they were under traditional testing. But we have made an enormous investment in faculty time and energy over the past decade to develop the system, and the ongoing corporate work of evaluating assessment — both the several instruments and the whole system — represents a new area of faculty responsibility.

Besides the quantity of effort, our exploration into assessment has also required a new quality or kind of work from us as a faculty. Setting aside our accustomed commitment to the most thorough possible discussion and academic debate, we had to commit ourselves to active experiment. We had to "agree to agree," to move forward and work with a liveable plan even though many of us might harbor reservations and uncertainties, in order to gain some concrete experience to learn from and improve upon. We had to deny ourselves the luxury of standing at an impasse, in the interest of making at least an honest attempt.

We feel certain, however, that the benefits have already outweighed the costs. The assessment system itself, in what it has already contributed to our students' learning, has proved beyond price. In addition, we have gained as a faculty. We are much more able than ever before to work openly and effectively as colleagues, without regard to status or discipline. We have come to rely so habitually on a constant, collaborative sharing that our unusually full schedule of workshops and institutes never completely satisfies the demand. We are learning how to fail, to admit failures and even to share them so they can be learned from. We are thus also gaining a confidence that gives our teaching more authority and more flexibility than our students used to be able to expect.

In writing, this all sounds much more clearly planned and organized than it ever was, is or will be in experience. Our exploration began as a college-wide effort because we were confronting a college-wide rethinking of our mission and our goals. It might just as easily have begun in a single department or division or professional program. Where we feel we have succeeded is in discovering — and learning some of the ways to maintain and nurture — a dynamic process of development.

At the heart of that process is the willingness to risk. Most of our learning has come from experience: we have never yet been able to foresee and predesign nearly as well as we can understand and modify afterward. The wisdom of hindsight is priceless, but there's only one way to get it. As one faculty member put it, we couldn't begin working on our future until we took the leap that brought it into our present.

Elements of Stability

The essential precondition for maintaining this coherence and quality control is **faculty consensus**. Without the months and years of open discussion and debate, experiment and reflection, brainstorming and critiquing that went into the initial defining of our renewed commitment to liberal education, any attempt to develop an assessment system would have been futile.

The original dialog between the faculty and its academic task force has taken more stable forms now. But the commitment to Friday afternoon workshops and regular faculty/staff institutes coordinated by the Faculty Development Committee remains, as a concrete way of maintaining the free flow of information and discussion we need to keep our consensus intact and vital. That commitment's depth is reflected in the intense competition each semester among faculty teams, departments and committees and support offices eager to share their work with their colleagues.

The most important single outcome of our faculty's hard-won and dynamic consensus is the **curriculum** itself. The curriculum, as an integrated matrix of developmental experiences, also provides the strongest framework of stability for our assessment system. Because the curriculum is by design coherent and unified — rather than a collection of programs designed for disparate goals and often working at cross-purposes — it could and did give rise to a coherent, unified assessment system.

This coherence and unity also make it possible for us to encourage students to pursue highly individual paths through the curriculum, since its developmental logic permeates every course and program. We have indeed managed to approach a situation in which, as one faculty member put it, "the student bumps into what she needs next whichever way she turns."

Working from consensus, we have been able to progress steadily but unhurriedly toward **generic assessment techniques**. These can provide an important unifying framework within an assessment system, but they could never have been successfully imposed nor could the process of their discovery be abridged without vitiating their meaning.

Collaborating on our common venture we have come to perceive the patterns underlying our several disciplines' approaches to a given competence, and have discerned basic unities within the diverse assessment techniques invented for a given level by various instructors. Gradually, with careful testing and faculty feedback for redesign, we have been able to offer more and more generic criteria and even complete instruments that can be used effectively — with appropriate adaptations and augmentations — throughout the college.

This work of monitoring assessment techniques for comparability of meaning and developing generic instruments and criteria is a major responsibility of the Assessment Committee. This body, with representatives from each Competence Division, the Assessment Center and related support areas, is charged with overseeing the entire assessment system as it evolves.

The work is also carried on simultaneously in several other areas. It is a continuing part of the collaborative work of each academic and professional department, and likewise of each of the Discipline Divisions into which the departments are organized. Each of the eight interdisciplinary Competence Divisions also has a major responsibility to review assessment techniques across disciplines (as outlined in Chapter Two above) and to operate as clearinghouse, coordinator and generic designer.

External assessment techniques also provide a valuable stabilizing element in a developmentally fluid system such as ours. As with the generic techniques, external assessments could not have been meaningfully imposed at the outset. They have evolved as the faculty grew more and more convinced that the externality inherent in all our assessments could and should be extracted, as it were, in a purer form to give us greater access to outside perspectives on a regular basis.

Assessors may be drawn from outside the course or learning experience or from off-campus. The entire assessment situation may be administered outside the course setting, as are those handled through the Assessment Center. Perhaps the most concentrated dose of externality comes in assessment techniques that are not designed as the culmination of any course or learning experience, such as the pre-entry assessment battery for level 1 of communications or the Integrated Competence Seminar at the end of the second year.

Wholly external assessments like the ICS, however, require great restraint. Such an assessment's value to the faculty as an outside "check" on what is being taught and learned must be weighed against its tendency to be of limited developmental value to students, who experience it as something for which they have not had adequate learning and with which they have no further opportunity to work once they have taken it. For this reason the initiation and design of the ICS and any similar instruments lies with the Assessment Committee.

As the Competence Divisions review their respective competences, the Assessment Committee regularly reviews the "whole picture" across departments and disciplines and among the several competences. It is their task to see that a potentially important development in one area is disseminated — even before a fuller report in a forum like the faculty/staff institute — into other areas

These represented qualities — integration, independence, self-awareness, creativity, commitment and habituality — that seem to stand in relation to competence as an adverb does to a verb. Thus, they reasoned, a student might demonstrate her ability to "take a personal position regarding the implications of contemporary events" (contemporary world, level 6) or to "lead organizational activity effectively" (social interaction, level 6) and be validated. But she might further be identified in terms of how self-aware she was about the techniques she used in her leadership, or to what extent she integrated her stands on particular issues within a coherent personal framework.

Using the results from small-group brainstorming sessions with the whole faculty and staff, the Assessment Committee returned to work on the characteristics to break out their modes of appearance at different levels of student development. Work on the "advanced characteristics" is still in its early stages — but clearly this new exploration will make a salient contribution to the future shape of our assessment system.

Nor is the Assessment Committee the only institutionalized agent for change. At every other level as well, from the Competence Divisions to the individual faculty member, the process of pulling together what has been learned inevitably suggests new questions and directions and the solution to one challenge almost always poses others.

This intertwined process can even be seen in the work of our Office of Evaluation. The office was created as a permanent means of enabling us to stand outside our own assumptions and continually critique the relevance of our educational goals and the effectiveness of our methods. Drawing on diverse instruments and techniques for assessing human development, cognitive and other skills and various modes of learning, the Office of Evaluation staff conducts a number of cross-sectional and longitudinal studies of our students' development according to standards and measures other than our own.

They also design and administer intensive interview techniques for eliciting student perceptions and attitudes during their years here, and maintain an ongoing program of comparison testing among assessment instruments from throughout the college. In addition, they are conducting studies of outstanding professionals in several fields and have planned follow-up studies of our graduates as means of checking the value of our curriculum and competences to the professional woman.

Tradition would suggest carrying on such research as far outside the college as possible. From the first, however, we realized that existing measures and experimental designs would need serious modification for use in evaluating our non-traditional methods. The Office of Evaluation staff would therefore need to have as thorough an understanding of our methods and their rationale as any faculty member. It also soon became clear that as our understandings and methods continued to evolve, the evaluation team would need to be intimately aware of the changes in order to keep their monitoring up to date.

As a result the Office of Evaluation has operated from its inception as an integral part of the college, and in collaborating with faculty members to design and revise their own studies the evaluation team has made innumerable contributions to the faculty's thinking and planning. As they are completed and published, the Office of Evaluation's reports will be as valid and objective as the best research skills can make them. Even while they are being carried out, however, the office's monitoring and evaluation efforts exert a strong and invaluable creative influence on our curriculum in general and on our assessment system in particular.

A Heuristic for Developing Systems

Because constant evolution of this sort is not only inevitable but desirable, and because each faculty's best effort to assess their achievement of their educational aims will necessarily result in a unique system, we have concluded that the most useful model we can offer is not a detailed chart of our own organization or the flows of process within it. Instead, we feel we can best share what we have learned by offering another heuristic — this time not a model, but a series of questions indicating dynamics or elements which we have found essential to the development and ongoing growth of our assessment system.